```
HCV capsid peptide
HCV capsid peptide
HCV capsid peptide
NON-A, non-B hepat
Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
Human hepatitis C
HUW core-envelope
HCV core-envelope
                                                                                                                                                                                                                                                                                                                                                                                             Prototype peptide
HCV core protein p
PT-NANB viral stru
Hepatitis C virus
Antigen pHCal01.
Antigen pHCal01.
HCV fragment 1 / I
                                                    Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
Non-A, Non-B Hepat
Non-A, Non-B Hepat
Hepatitis C virus
CN14 fragment of H
                                                                                                                                                                                                                                                                                                                                                                    Peptide VIIIE base
Anti-HCV antibody
             NC mosaic protein
Blood transmiscibl
HCV-S1 full-length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                fragment 1 /
fragment 2 /
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NC mosaic protein amino acid fragment A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (USSH ) US DEPT HEALTH & HUMAN SERVICES
                                                    AAR84559
AAR74562
AAR74562
AAR24946
AAR22956
AAR32056
AAR30689
                                                                                                                                                                                                                                                                               AAR29534
AAR29535
AAR29536
AAR29537
                                                                                                                                                                                                                                                                                                                                       AAR29542
AAR25121
AAR20770
AAR69545
AAY15425
AAW11865
AAR12597
AAR96530
AAR29160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR41422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR41423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY06673 standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17-JUN-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Khudyakov YE;
                                                    470
28
3010
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1999-204671/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO9910506-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04-MAR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Fields HA,
AAY06673;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY0667.
HCV type 3 capsid
Hepatitis C virus
Encoded by Hepatit
Encoded by Hepatit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           mosaic protein
                                                                                          Search time 38.5455 Seconds (without alignments)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                   1. SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2. SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3. SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
4. SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
5. SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
5. SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
6. SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
7. SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
8. SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
8. SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
8. SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
8. SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
8. SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
8. SIDSI/gcgdata/geneseqy-embl/AA1982.DAT:*
8. SIDSI/gcgdata/geneseqy-embl/AA1992.DAT:*
8. SIDSI/gcgdata/geneseqy-embl/AA1992.DAT:*
8. SIDSI/gcgdata/geneseqy-embl/AA1993.DAT:*
8. SIDSI/gcgdata/geneseqy-geneseqy-embl/AA1993.DAT:*
8. SIDSI/gcgdata/geneseqy-geneseqy-embl/AA1993.DAT:*
8. SIDSI/gcgdata/geneseqy-geneseqy-embl/AA1993.DAT:*
8. SIDSI/gcgdata/geneseqy-geneseqy-embl/AA2003.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                               1107863
          GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                             al number of hits satisfying chosen parameters:
                                                                                                                                                                151
1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                  1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                             August 7, 2003, 11:05:37;
                                                                                                                                                                                                                                                                                                                                                   Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB71258
AAR92968
AAR92969
AAR92971
AAR96547
AAR92970
AAR34473

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY06673
                                                                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                            A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ü
                                                                                                                                                                                                                                                                                                           length: 0
length: 2000000000
                                                                                                                                                   US-09-491-146A-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20
23
23
11
11
14
14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28
189
191
191
191
191
470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100.0
100.0
100.0
100.0
100.0
99.3
97.4
```

Score

Result Š 151 151 151 151 151 151 147

sed

Minimum DB Maximum DB

Database

Searched:

Perfect score: Sequence: Scoring table:

protein

: 0

Run 8

Gaps

ö

Indels

 $\overset{\times}{\times}$   $\overset{\times}$ 

ò a

```
wide range of flavivirus diseases, e.g. yellow fever, Japanese encephalitis, dengue, classical swine fever, bovine viral diarrhoea and hepatitis C. The specified deletion: (i) produces a reliably attenuated virus that does not revert to virulence; (ii) is exactly defined and does not effect immune responses to important proteins; and (iii) can not generate a non-natural virus by recombination. The mutant viruses eliminate the need to produce large amounts of infectious/virulent viruses, and can be produced with less expense. The protective response to flavivirus lasts significantly longer than that to killed vaccines. This sequence represents a fragment of the capsid protein from Hepatitis c virus (HCV) type I described in the disclosure of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100.0%; Score 151; DB 17;
100.0%; Pred. No. 1.1e-13;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                             ; Score 151; DB 23;
; Pred. No. 1.1e-13;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus isolate HK10.core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 4 PKPQKKTKKNTIRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEPT HEALTH & HUMAN SERVICES SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR92968 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 4; Page 207; 340pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Purcell RH;
                                                                                                                                                                                                                                                                                                                               100.0%;
100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             02-OCT-1996 (first entry)
                                                                                                                                                                                                                                                                                                                Ouery Match
Best Local Similarity 100.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1996-139709/14.
                                                                                                                                                                                                                                                                                 189 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAT16642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9605315-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22-FEB-1996.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SD (HSSD)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hepatitis.
                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR92968;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR92968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                              The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The artigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This invention describes a novel attenuated flavivirus live vaccine comprising a flavivirus mutant that has a deletion of at least 4 consecutive amino acids from the capsid protein, provided that the C-terminal hydrophobic region is not affected by the deletion. The vaccine of the invention has virucide, antinflammatory and hepatotropic activity. The attenuated vaccine, and similar nucleic acid vaccines that encode the mutated capsid protein, are useful for protection against a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Attenuated flavivirus live vaccine, useful for protection against e.g. yellow fever, comprises virus with attenuating deletion of amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Capsid protein; attenuated vaccine; virucide; antinflammatory; hepatotropic; yellow fever; Japanese encephalitis; dengue; classical swine fever; bovine viral diarrhoea; hepatitis C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                               100.0%; Score 151; DB 20;
100.0%; Pred. No. 1.5e-14;
iive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV type 3 capsid protein fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB71258 standard; protein; 189 AA
                                                                                                                                                                                                                                                                                                                                                                        comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; Fig 2; 30pp; German.
                        Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11-FEB-2002; 2002WO-AT00046.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21-FEB-2001; 2001AT-0000272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18-NOV-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match est Local Similarity luv...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      from the capsid protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2002-667064/71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Heinz FX, Mandl C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (HEIN/) HEINZ F X. (MAND/) MANDL C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis c virus
                                                                                                                                                                                                                                                                                                                                                                                                                      28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200266621-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29-AUG-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB71258;
                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (MAND/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB71258
```

ö

Gaps

ô

Indels

Length 191;

ò g us-09-491-146a-23.rag

```
/note= "amino acid in this position is designated X in the specification, but codon usage shows that the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus; subtype; polymerase chain reaction; amplification;
PCR; primer; probe; antibody; infection.
                                                                                                                                                                                                                                                                                                                                                                                              AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, sallva, lymphocytes or other monounclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                   DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vactines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus types 7c(8a) isolates VN4 amino acids 1-317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 151; DB 17;
Pred. No. 1.1e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 144..149 / /label - Val, Ala, Glu, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /label- Met, Thr, Lys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US DEPT HEALTH & HUMAN SERVICES. US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /label- Val, Ala, Asp, Misc-difference 161
                                                                                                                                                                                                                                                                                                                                                                  Claim 4; Page 209-210; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /label= Met, Leu
                                                                                                                                                                                                                                        Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR96547 standard; peptide; 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 100.0%;
Best Local Similarity 100.0%;
Matches 28; Conservative 0
                                                                                                                         95WO-US10398
                                                                                                                                                          94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /label- Gly
                                                                                                                                                                                                                                        Bukh J, Miller RH,
                                                                                                                                                                                                                                                                      WPI; 1996-139709/14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 144
                            Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
                                                                                                                                                                                                                                                                                       N-PSDB; AAT16645.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              191
                                                             WO9605315-A2
                                                                                                                           15-AUG-1995;
                                                                                                                                                        15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10-MAR-1997
                                                                                           22-FEB-1996
hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR96547;
                                                                                                                                                                                        ( HSSU)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR96547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                       envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               proteins -
HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA and amino acid sequence of HCV envelope 1 and core used to determine HCV genotype and as vaccines against
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus isolate DK12 core protein.
                                                                                                                                                                                                                        Hepatitis C virus isolate S52 core protein.
                               PKPQKKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                               (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR92971 standard; Protein; 191 AA.
                                                                                                                           AAR92969 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 4; Page 208; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bukh J, Miller RH, Purcell
                                                                                                                                                                                                                                                                                                                                                                                                   95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                  94US-0290665
                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1996-139709/14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       191 AA;
                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-PSDB; AAT16643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                   15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                  15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                    WO9605315-A2
                                                                                                                                                                                        02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                  22-FEB-1996,
                                                                                                                                                                                                                                                                      hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR92971;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                          AAR92969;
                                                                                                                                                                                                                                                      HCV; E1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR92971
ID AAR
                                                                                                             AAR92969
                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EXEX DX XX
```

x S

à

```
HCV; E1; envelope 1; core protein; HCV genctyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, sellva, lymphocytes or other monounclear coils. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                / Match 99.3%; Score 150; DB 17; Length 191; Local Similarity 96.4%; Pred. No. 1.6e-13; les 27; Conservative 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV; non-A, non-B hepatitis virus; NANBHV; liver disease; polymerase chain reaction; diagnostic method.
                                                                                                                                            Hepatitis C virus isolate S2 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Encoded by Hepatitis C virus clone JK3-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (USSH ) US DEPT HEALTH & HUMAN SERVICES (USSH ) US SEC DEPT HEALTH.
                     AAR92970 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR34473 standard; Protein; 470 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 4; Page 209; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                          95WO-US1039B
                                                                                                                                                                                                                                                                                                                                                                                                              94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1996-139709/14.
                                                                                                                                                                                                                                                Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAT16644.
                                                                                                                                                                                                                                                                                         WO9605315-A2,
                                                                                                                                                                                                                                                                                                                                                                                                              15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                      15-AUG-1995;
                                                                                                        02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30-JUL-1993
                                                                                                                                                                                                                                                                                                                               22-FEB-1996
                                                                                                                                                                                                          hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                             AAR92970;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR34473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
    AAR92970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR34473
                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The sequences AAR96526-R96578 represent novel sequences isolated from hepatitis C virus subtypes different from subtypes la-c, 2a-d, 3a-f, 4a-j, 5a and 6a. They esp, from the novel subtypes la-f, 2e-i, 2k, 2l, 3q, 4k-m, 7a-c or types 9, 10 or 11. The sequences corresp. to the 5' untranslated region (UR), the Core/El, NS4 or NS5B regions of the genome. This sequence represents amino acids 1-317 from the HCV types 7c and 8a isolates VM4.

The new HCV types were isolated from patients with chronic HCV from the Benelux countries, France, Cameroon and Vietnam, because of their abenelux countries. France, Cameroon and Vietnam, because of their abenelux countries. France, Cameroon and Vietnam, because of their applified, cloned and genotyped. The 5'UR, Core/El and NSSB regions were sequenced aither directly or patielally and used to classify the new viruses into (sub)types based on comparison with known sequences. The sequences can also be used to synthesise probes and primers for the sequences can also be used to synthesise probes and primers for the detection of HCV in a sample. The polypeptides can be used to detect
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                        /note= "amino acid in this position is designated X in the specification, but codon usage shows that the only possible amino acid at this pos. is Gly"
only possible amino acid at this pos. is Gly"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus poly:nucleic acid unique to unidentified sub:type - used to develop probes and primers for new sub:types and vaccines to prevent and treat infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0%; Score 151; DB 17;
100.0%; Pred. No. 1.9e-13;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                 Leu, Ile, Val
                                                                              'label≖ Val, Ala, Glu, Gly
                                                                                                                                                                                                                         label- Cys, Arg, Ser, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                        label- Ser, Arg
                                                                                                                                                                                                                                                                                                          Len
                                                                                                                                                                                                                                                                                                                                                Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 25; Fig 3; 150pp; English.
                                                                                                                                                                                                                                                                                                                                    /label- Met,
Misc-difference 233
                                                                                                                                                                                                                                                                 'label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                         /label- Asn,
                                                                                                                                                                                                                                                                                                        'label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                95WO-EP04155.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     95EP-0870076.
                                                                                                                   /label- Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (INNO-) INNOGENETICS NV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Maertens G, Stuyver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1996-251460/25.
N-PSDB; AAT27958.
                   Misc-difference 167
                                                                                                                                                                                                      Misc-difference 172
                                                         Misc-difference 167
                                                                                                                                                                                                                                            Misc-difference 174
                                                                                                                                                                                                                                                                                                                         Misc-difference 232
                                                                                               Misc-difference 17
                                                                                                                                                                                                                                                                                   Misc-difference 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 319 AA;
```

28-JUN-1995; 21-OCT-1994; 23-OCT-1995;

WO9613590-A2 09-MAY-1996.

ö

Gaps

ö

JP05068562-A.

23-MAR-1993

RESULT

ò g

38888888888**%**&

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JX3-C. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV.
                               Gaps
                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 470;
 Length 470;
                                                                                                                                                                                                                                                                                                        HCV; non-A, non-B hepatitis virus; NANBHV; liver disease; polymerase chain reaction; diagnostic method.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA and cDNA of hepatitis C virus – useful as probes for diagnosing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97.4%; Score 147; DB 14;
llarity 96.4%; Pred. No. 1.1e-12;
Conservative 1; Mismatches 0;
Score 147; DB 14;
Pred. No. 1.1e-12;
1; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Encoded by Hepatitis C virus clone JK3-C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NC mosaic protein amino acid fragment C.
                                                               28
                                                                              PKPQRQTKRNTIRRPQDVKFPGGGQIVG 32
                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 4; Page 30-32; 44pp; Japanese.
                                                                                                                                                                             AAR34475 standard; Protein; 470 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (SANW ) SANWA KAGAKU KENKYUSHO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28
Query Match 97.4%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY06675 standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1993-130638/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            See AAQ40425-Q40439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
nes 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAQ40433
                                                                                                                                                                                                                                                                                                                                                                                      JP05068562-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      17-JUN-1999
                                                                                                                                                                                                                                          30-JUL-1993
                                                                                                                                                                                                                                                                                                                                                                                                                     23-MAR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY06675;
                                                                                                                                                                                                            AAR34475;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY 06675
                                                                                                                                                              AAR34475
                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ó
                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-B. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV. See AAQ40425-Q40439.
                                                                                                                                                                                                                                        cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-A. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV.
                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                          Length 470;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; non-A, non-B hepatitis virus; NANBHV; liver disease; polymerase chain reaction; diagnostic method.
                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA and cDNA of hepatitis C virus - useful as probes for
                                                                                                                                                              DNA and cDNA of hepatitis C virus - useful as probes for
                                                                                                                                                                                                                                                                                                                                                                        Score 147; DB 14;
Pred. No. 1.1e-12;
1; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                   Encoded by Hepatitis C virus clone JK3-B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 4; Page 28-30; 44pp; Japanese.
                                                                                                                                                                                                         Claim 4; Page 26-28; 44pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR34474 standard; Protein; 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (SANW ) SANWA KAGAKU KENKYUSHO
                                                                              (SANW ) SANWA KAGAKU KENKYUSHO
                                                                                                                                                                                                                                                                                                                                                                       Match 97.4%;
Local Similarity 96.4%;
les 27; Conservative 1
                  91JP-0153736.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91JP-0153736.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91JP-0153736
                                                91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       diagnosing HCV infection
                                                                                                                                                                                diagnosing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1993-130638/16.
N-PSDB; AAQ40432.
                                                                                                               WPI; 1993-130638/16.
                                                                                                                                                                                                                                                                                                             See AAQ40425-Q40439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     470 AA;
                                                                                                                                                                                                                                                                                                                                           470 AA;
                                                                                                                                 N-PSDB; AAQ40431
                  30-MAY-1991;
                                                30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30-JUL-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JP05068562-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             23-MAR-1993
                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR34474;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR3447
```

ž

ò a

```
Gly
                                                                                                                                                                               Leu
                                                                                           /label= Ser, Arg,
                                                                                                                                                                             label- Gln, Pro,
                                                     label = Gly, Asp
                                                                                                                                                                                                                   label- Lys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               abel- Ala, Thr
/label- Phe,
                                                                                                                                                                                                                                                                                                  .abel- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                            abel- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         abel- Thr,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            label- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                label- Gly,
                                                                                                                                    'label Thr,
                                                                                                                                                                                                                                                          abel- Gly,
                                                                                                                                                                                                                                                                                                                                           abel- Phe,
                                                                                                                                                                                                                                                                                                                                                                                   label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  label- Glu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          label- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       label= Glu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label Lys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             'label- Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                'label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       'label= Ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92JP-0051885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92JP-0051885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 1036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 2462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 2488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 2989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 1201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 1828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 2143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 2144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 1056
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 1255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 126
                                                                                                               Misc-difference 400
                                                                                                                                                     Misc-difference 405
                                                                                                                                                                                              Misc-difference 410
                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 759
                                                                      Misc-difference 398
                                                                                                                                                                                                                                                                           Misc-difference 430
                                                                                                                                                                                                                                                                                                                   Misc-difference 43
                                                                                                                                                                                                                                                                                                                                                            Misc-difference 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                             Misc-difference
                                                                                                                                                                                                                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JP06105690-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .0-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-APR-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences ANYOGGS12681 represent amino acid sequence of each monomer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polymerase chain reaction; PCR; amplify; primer; non-A, non-B hepatitis; NANBH; virus; blood transmissible; detection; hepatitis virus; RT-PCR; C100 antibody; HCV RNA; NSS region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                           New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    96.7%; Score 146; DB 20;
96.4%; Pred. No. 7.9e-14;
Live 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                     (USSH ) US DEPT HEALTH & HUMAN SERVICES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR53417 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Blood transmiscible NANBHV protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              comprising the NC mosaic protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /label- Asp, Asn
Misc-difference 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /label= His, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          'label = Cys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /label- Leu, Phe
                                                                                                                                                                                                                                                                                                                                                                                              Claim 5; Fig 9; 66pp; English,
                                                                                                        98WO-US17385.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Non-A, non-B hepatitis virus
                                                                                                                                                 97us-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  17-JAN-1995 (first entry)
                                                                                                                                                                                                                              Fields HA, Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 96.4 les 27; Conservative
                                                                                                                                                                                                                                                                  WPI; 1999-204671/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28 AA;
                         WO9910506-A1
                                                                                                                                                 25-AUG-1997;
                                                                                                        21-AUG-1998;
                                                                04-MAR-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR53417;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 12
```

ò g ò

Gaps

ö

Indels

Length 3010;

```
N-terminal portion of a coding sequence of RNA virus, a second polynucleotide region including a 3' untranslated region (UTR) sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus and a third polynucleotide region encoding a reporter molecule, flanked by first and second polynucleotide regions; and a promoter sequence being operatively linked to expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette. Nucleic acid construct of the invention is useful for detecting the presence of an RNA virus in a cell. It is also useful for screening anti-viral drugs and determining drug resistance of an RNA virus. The present sequence is Hepatitis C virus (HCV) isolate HCV-SI full-length polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The present sequence is that of an antiqen used in a novel assay determining anti-HCV antibody titre. A sample is diluted to an absorbance of not more than 2.0 and the peptide antigen is added. The method is useful for monitoring the effect of interferon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Determination of the antibody titre against hepatitis C virus antigen - using dilute solution and HCV antigenic peptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV; non-A, non-B hepatitis virus; antigen; immunoassay; interferon treatment; monitoring; antibody titre; detection
                                                                                                                                                                                                                                                                                                                 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 145; DB 16;
Pred. No. 1.6e-13;
                                                                                                                                                                                                                                                                                                               Score 146; DB 2
Pred. No. 1e-11;
                                                                                                                                                                                                                                                                                                                                                       0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus core antigenic peptide.
                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  treatment on anti-HCV antibody titres.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 1; Page 10; 11pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             96.0%;
96.4%;
                                                                                                                                                                                                                                                                                                               96.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    94JP-0082160.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94JP-0082160.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR84559 standard; peptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (ARIM/) ARIMA T.
(TORA ) TORAY IND INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1995-386063/50.
                                                                                                                                                                                                                                                                          3010 AA;
                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
les 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        40 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JP07260792-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              06-JUN-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13-OCT-1995
                                                                                                                                                                                                                                                                          Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR84559;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR84559
    8$66666666668$
                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                         This sequence is encoded by the genome of a blood transmissible non-A, non-B hepatitis (NANBH) virus. The CDNA sequence was isolated using the primers given in AAQ63500-35. The amplified fragments are used in the detection of hepatitis virus. The target DNA was isolated from serum of chronically infected NANBH patients who were Cl00 antibody-positive were performed on cDNA and the total human NANBH DNA was constructed from 23 clones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Novel nucleic acid construct useful for detecting the presence of RNA virus, comprises an expression cassette and a promoter operably linked to expression cassette for minus strand RNA transcription of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The invention relates to nucleic acid construct which comprises an expression cassette including a first polynucleotide region including a 5' non-coding region (NCR) sequence of an RNA virus and at least an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nucleic acid construct; expression cassette; non-coding region; NCR; untranslated region; UTR; anti-viral drug; drug resistance;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                            Blood-transmissible non-A non-B hepatitis virus DNA - used for
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                      / Match 96.7%; Score 146; DB 15;
Local Similarity 96.4%; Pred, No. 1e-11;
nes 27; Conservative 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (MOLE-) INST MOLECULAR & CELL BIOLOGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Example 1; Page 70-81; 81pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ¥
                                                                                                                                                                                        Claim 1; Page 8-20; 22pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAE20477 standard; Protein; 3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV-S1 full-length polyprotein
                                                                                                                                                  detection of hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20-JUL-2001; 2001WO-IL00669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24-JUL-2000; 2000US-220248P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV-S1; Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ran YH, Lim SP, Lim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 2002-280605/32.
                                                              WPI; 1994-163130/20.
                                                                                                                                                                                                                                                                                                                                                                                                                 3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EHRL/) EHRLICH G
                       (KAEN/) KAENNO K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; AAD33038
                                                                                     N-PSDB; AAQ63499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200208447-A2
```

01-JUL-2002

AAE20477;

RESULT 13 AAE2047

a

Sequence

Query Match

Matches

31-JAN-2002.

ö

Gaps

ö

Length 40; Indels

assay for

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR74561-63 and AAR74566 are respectively, 24mer, 40mer, 50mer and 30mer antigenic peptides derived from the hepatitis C virus clone S29 genes III and IV. They are useful in the detection of and diagnosis of Hepatitis C virus infection. The peptides are used in a detection method which has a higher sensitivity than conventional detection agents.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                              Agent for detection of hepatitis C - also hepatitis C virus (HCV) core protein antigen peptide(s), useful for immuno:detection of HCV \,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            96.0%; Score 145; DB 16; Length 40; 96.4%; Pred .. No. 1.6e-13; Live 0; Mismatches 1; Indels
                                                                                                                                   Hepatitis C virus; detection; immunodetection; sensitive.
                                                                                                          Hepatitis C virus core protein antigen peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4 PKPQRKTKRNTSRRPQDVKFPGGGQIVG 31
                        AAR74562 standard; peptide; 40 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 1; Page 7; 10pp; Japanese.
                                                                                                                                                                                                                                                   93JP-0273938.
                                                                                                                                                                                                                                                                               93JP-027393B
                                                                          04-JAN-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 96.4 Matches 27; Conservative
                                                                                                                                                                                                                                                                                                         (ARIM/) ARIMA T.
(TORA ) TORAY IND INC.
                                                                                                                                                                                                                                                                                                                                                  WPI; 1995-175362/23.
                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  40 AA;
                                                                                                                                                                                           JP07097398-A.
                                                                                                                                                                                                                                                   28-SEP-1993;
                                                                                                                                                                                                                                                                             28-SEP-1993;
                                                                                                                                                                                                                       11-APR-1995.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                    AAR74562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
RESULT 15
                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          돈
```

>earch completed: August 7, 2003, 11:14:05
Job time : 39.6364 secs

polyprotein polyprotein

genome

genome polyprotein genome polyprotein polypeptide - hepa

hypothetical prote

genome polyprotein genome polyprotein

genome polyprotein

genome polyprotein genome polyprotein genome polyprotein

tein - hepa polyprotein polyprotein

polyprotein polyprotein genome polypr

```
Genome polyprotein - hepatitis C virus (genotype 3, NI) (fragment)
N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 3, NI
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C.Accession: $41358
R.Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A.Reference number: $41341
A.Accession: $41358
A.Molecule type: genomic RNA
A.Residues: 1-114 < VANA
A.Residues: 1-114 < VANA
A.Residues: 1-114 < VANA
A.Cross references: EMBL: 229461; NID: 9443884; PIDN: CAA82599.1; PID: 9443885
A.Gross references: Embl. C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
F:1-114/Product: core protein #status predicted < CMAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                 Userious programs. Applications of the protein of t
                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (genotype 3, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 100.0%; Score 151; DB 2; Best Local Similarity 100.0%; Pred. No. 2.4e-14; Matches 28; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                              ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 151;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                          $18031
$18032
PN0677
JQ0883
JQ0881
                                   PC1284
JQ1925
JQ1926
                                                                                        JH0711
S19876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.08;
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 2
 polyprotein polyprotein polyprotein polyprotein polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          yprotein
yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Lyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein
                                                                                                                        August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    pol y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       pol
pol
pol
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              po]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome
genome
genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome
genome
genome
genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         зепоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              зепопе
                GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                 .al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                              283308 seqs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                          protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A45573
S21336
S21335
S41355
S41355
S41357
S41341
S41350
S41350
S41354
S41354
S41354
S41354
S41354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S41358
S41360
PC2061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             S41288
S41361
S19875
S41356
PC2060
                                                                                                                                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  541343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S41359
                                                                                                                                                                                                                                                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                   US-09-491-146A-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             492
123
782
108
411
3010
88
108
108
108
112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
1: pir4:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      96.0
95.4
95.4
95.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100.0
100.0
97.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  96.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query
                                                                                                                                                                                                                Perfect score:
Sequence:
                                                                                                                                                                                                                                                                         Scoring table:
                                                                                          protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Database :
                                                                                                                                                                                                                                                                                                                               searched:
                                                                                                                              Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ş
```

ö

Gaps

ö

Indels

Length 114;

DB 2;

Length 114;

Gaps

```
Syseligy, R.
submitted to the EMBL Data Library, December 1993
A; Reference number: $41288
A; Accession: $41288
A; Accession: $41288
A; Molecute type: genomic RNA
A; Residues: 1-492 <SEE>
A; Cross-references: EMBL:X76918
A; Cross-references: EMBL:X76918
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein: core protein; envelope protein; nonstructural protein; F; 1-191/Product: core protein #status predicted <CNN>
F; 192-372/Product: NS1 protein (fragment) #status predicted <NS1>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Accession: S19875
A;Molecule type: genomic RNA
A;Residues: 1-782 <-(HON>
A;Coss-references: EMBL:X61592; NID:959482; PIDN:CAA43789.1; PID:959483
A;Experimental source: isolate JK3
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstruct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (isolate JK3) (fragment)
N;Contains: core protein; envelope protein 1; nonstructural protein 2; NSI/E2 prote
C;Species: hepatitis C virus
          C;Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000 C;Accession: S41288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C. Accession: S41361
R. Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A. Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A. Reference number: S41341
A. Accession: S41361
A. A. Residues: 1-123 cVAN>
A. Residues: 1-123 cVAN>
A. Residues: 1-123 cVAN>
A. Residues: 1-123 cVAN>
A. Experimental source: genoetype 3, NA
A. Experimental source: genoetype 3, NA
Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Variety: isolate JK3
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C.Accession: S19875
R.Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A.Description: Sequence analysis of putative structural regions of A.Accession: S18029
A.Accession: S19875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatitis C virus (genotype 3, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 492;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match 100.0%; Score 151; DB 2; Best Local Similarity 100.0%; Pred. No. 9.9e-14; Matches 28; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97.4%; Score 147; DB 2; 92.9%; Pred. No. 9.5e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Variety: genotype 3, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N; Contains: core protein
C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S19875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               olecule type: mRNA

v. kesidues: 1-411 < LLD>

A: Cross-references: 1-411 < LLD>

A: Cross-references: GB:L12355; NID:g410169; PIDN:AAA20155.1; PID:g410170

C: Superfamily: hepatitis c virus genome polyprotein

F: Stewwords: ATP: capsid protein is a *status predicted < SPE-

F: 192-383/Product: nonstructural protein E2/WS1 *status predicted < NPE-

F: 196, 209, 234, 305, 325/Binding site: carbohydrate (Asn) (covalent) *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genome polyprotein N2 - hepatitis C virus
N;Contains: envelope protein E1; nonstructural protein E2/NS1
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Accession: PC2061
R;Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
Biochem. Biophys. Res. Commun. 199, 1474-1481, 1994
A;Title: Identification of the third major genotype of hepatitis C virus in France.
Tcession: PC2061
                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus (genotype 3, N3) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 3, N3
C;Spate: 19-Way-1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-2000
C;Accession: S4,360
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the RMBL Data Library, January 1994.
A;Description: Analysis of hepatitis C virus genotypes I to 5 by LipA.
A;Reference number: 841341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A,Cross-references: EMBL:229463; NID:9443889; PIDN:CAA82601.1; PID:9443889
A,Experimental source: genotype 3, N3
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; corre protein; polyprotein
F;1-124/Product: core protein #status predicted <MAT>
                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 411;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 124;
                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein - hepatitis C virus (fragment)
N;Contains: core protein; envelope protein; NSI protein
C;Species: hepatitis C virus
   Pred. No. 2.4e-14;
Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100.0%; Score 151; DB 2;
100.0%; Pred. No. 2.6e-14;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. 8.3e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0%; Score 151;
100.0%; Pred. No. 8
:ive 0; Mismatche
                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                          Best Local Similarity 100.0%; F
Matches 28; Conservative 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 100.(
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Olecule type: genomic RNA .esidues: 1-124 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Accession: S41360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
S41288
                                                                                                     ô
                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
```

Hepatitis C Viru

ö

Gaps

m

Page

carr

(nons

à

```
genome polyprotein S29 (core protein region) - hepatitis C virus (fragment)
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 17-Nov-2000
C;Accession: S21336
A;Description: A sentitive serodiagnosis of hepatitis C virus infection with two c A;Reference number: S21336
A;Accession: S21336
A;Accession: S21336
A;Molecule type: genomic RNA
A;Residues: 1-88 (SAT>
                                                                                                                                                                                                                                                   from a single Japanese
                                           genome polyprotein - hepatitis C virus (strain JT)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98)
protein NS4s; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NID: 959492; PIDN: CAA46517.1; PID: 959493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (genotype 2, N2) (fragment) N;Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 88,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                           R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, Virus Res. 23, 39-53, 1992
Virus Res. 23, 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from A;Reference number: A45573; MUID:92295714; PMID:1318627
A;Accession: A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Match 96.7%; Score 146; DB 1; Local Similarity 96.4%; Pred. No. 3.1e-12; les 27; Conservative 0; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 145; DB 2;
Pred. No. 1.3e-13;
0; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKPORKTKRNTYRRPODVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKPORKTKRNTSRRPQDVKFPGGGGIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              y Match 96.0%;
Local Similarity 96.4%;
hes 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Cross-references: EMBL:X65548;
C;Superfamily: hepatitis C virus
C;Keywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S41353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 procession: protein N1 - hepatitis C virus

Gondafins: envelope protein E1; nonstructural protein E2/NS1

Species: hepatitis C virus

Contains: envelope protein E1; nonstructural protein E2/NS1

Species: hepatitis C virus

Catesion: PC2060

C; Accession: PC2060

R; Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.

Bacchem. Biophys. Res. Commun. 199, 14/4-1481, 1994

A; Title: Identification of the third major genotype of hepatitis C virus in France.

A; Recence number: PC2060

A; Molecule type: mRNA

A; Residues: 1-411 < LiJ>
C; Superfamily: hepatitis C virus genome polyprotein; glycoprotein; nonstructural protein; C; Superfamily: hepatitis C virus genome polyprotein

C; Seywords: APP: capsid protein: envelope protein; glycoprotein; G; Reywords: APP: capsid protein E1 ** status predicted <SPE>

F; 192-383/Product: envelope protein E1 ** status predicted <SPE>

F; 192-383/Product: nonstructural protein E2/NS1 ** status predicted

F; 196, 209, 234, 305, 325/Binding site: carbohydrate (Asn) (covalent) ** status predicted
                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
F;1-191/Product: core protein #status predicted <MAT1>
F;192-383/Product: envelope protein 1 #status predicted <MAT2>
F;384-733/Product: NS1/E2 protein #status predicted <MAT3>
F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>
                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (genotype 2, N5) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
'ariety: genotype 2, N5
ate: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41356
R;van Doorn, L.J.: Rleter, G.E.M.; Brouwer, J.T.
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Description: A1341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                        Score 147; DB 2; Length 782;
Pred. No. 5.8e-13;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 146; DB 2; Length 411;
Pred. No. 4.38-13;
2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A)Cross-references: EWBL:229459
A)Experimental source: genotype 2, N5
C)Superfamily: hepatitis C virus genome polyprotein
C)Keywords: capsid protein; core protein; polyprotein
F)1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 146; DB 2;
Pred. No. 1.2e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKPOROTKRNTIRRPODVKFPGGGGIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                        97.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  96.78;
92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          / Match 96.7%;
Local Similarity 96.4%;
les 27; Conservative
                                                                                                                                                    Best Local Similarity 96.4 Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Accession: S41356
A;Molecule type: genomic RNA
A;Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local S
Matches 27,
                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
```

ö

```
N; Contains:
                                                                                                                                                                                                                                                                                                                                                RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                     genome
                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
C; Species: hepatitis C virus
A; Variety: genotype 2, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41353
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41341
A; Accession: S41353
A; Molecule type: genomic RNA
A; Residues: 1-108 < VAN>
A; Residues: 1-108 < VAN>
A; Reperimental source: genotype 2, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted < MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       941357
genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
Nycontains: core protein
Nycontains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N6
C;Actesion: S41357
R;Van Doorn, L.J.: Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Reference number: S41357
A;Accession: S41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 2, N4
C.Date: 19-May 1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C.Accession: $41355
R.yam Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A.Reference number: $41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                           95.4%; Score 144; DB 2; Length 108; 96.4%; Pred. No. 2.2e-13; Live 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A: Experimental source: genotype 2, N4 C: Superfamily: hepatitis C virus genome polyprotein C: Keywords: capsid protein; core protein; polyprotein F:1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 144; DB 2; I
Pred. No. 2.2e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Cross-references: EMBL:229460
A;Experimental source: genotype 2, N6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
A; Cross-references: EMBL: 229458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 96.4
tches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Molecule type: genomic RNA Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Accession: S41355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
```

```
ö
                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                      C;Species: hepatitis C virus
A;Variety: genotype 1, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C; Accession: $41348
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPa.
A; Reference number: $41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                          - hepatitis C virus (genotype 1, N6) (fragment)
                                                                                                    Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: capsid protein; core protein; polyprotein F;1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A. Experimental source: EMBL:229451
A. Experimental source: genotype 1, N6
C. Superfamily: hepatitis C virus genome polyprotein
C. Keywords: cappid protein; core protein; polyprotein
F. 1-108/Product: core protein #status predicted <AMT>
                                                                                                    Score 144; DB 2;
Pred. No. 2.2e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 95.4%; Score 144; DB 2; Best Local Similarity 96.4%; Pred. No. 2.2e-13; Matches 27; Conservative 0; Mismatches 1.
                                                                                                                                                                                                7, 2003, 11:21:46
                                                                                                 Query Match 95.4%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: genomic RNA A; Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Search completed: August
Job time : 9.54545 secs
                                                                                                                                                                                                                                                                                                                                                                                                 core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: S41348
                                                                                                                                                                                                                                                                                                                                                                          polyprotein
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

US-09-491-146A-23 151 1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28

Scoring table:

127863 seqs, 47026705 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched:

127863 al number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

		Description	269	hepatitis						h genome	h genome	р депоше	c	h genome	P26664 h genome po	P10723 brugia mala		P19711 bovine vira	080361 nicotiana t			-		fugn		mus m	015530 homo sapien	Q922a0 mus musculu	O55173 rattus norv	Q28193 bos taurus	P51816 homo sapien		P31376 saccharomyc	P10637 mus musculu	P19332 rattus norv
SUMMARIES		QI	POLG_HCVJT	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVTW	POLG_HCV1	SYN_BRUMA	FURI_HUMAN	POLG_BVDVŅ	RK4_TOBAC	CAT3_MAIZE	FN3X_HUMAN	Y401_BUCAP	RL29_YEAST	RS7_FUGRU	RS7_HUMAN	FURI_MOUSE	PDPK_HUMAN	PDPK_MOUSE	PDPK_RAT	FURI_BOVIN	FMR2_HUMAN	YB1_XENLA	YAB1_YEAST	TAU_MOUSE	TAU_RAT
		8	7	7	-	-	-	-	ч	-	-		-	-	-	-	-	-	<del>, -4</del>	-	-	7	-	_	<b>-</b>	-	-	-	-	-	-	-		-	7
		Length	e	513	520	520	737	737	3010	3010	3011	3033	3033	3010	3011	548	794	3988	282	496	309	219	28	194	194	793	256	559	559	797	1311	303	638	732	751
ď	Query	Match Length	96.7	95.4	95.4	95.4	95.4	95.4					ī	$\sim$	90.1			m		32.5	$\sim$	_	31.1	-	31.1	ä	ö		ö	30.5			30.1	30.1	30.1
		Score	146					144	144			144	144	137	136	54	52	50.5	49	64	48.5	48	47	47	47	47	46	46	46	46	46	ഗ	45.5	വ	45.5
	Result	No.	ч	7	m	4	2	9	7	80	σ.	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	56	30	31	32	33

Q9z0e2 mus musculu	049937 spinacia ol	P37271 arabidopsis	P21525 drosophila	091713 xenopus lae	095405 homo sapien	P47390 mycoplasma	027036 methanobact	P08855 oryctolagus	P48155 manduca sex	P02362 xenopus lae	Q9czx5 mus musculu
CHRD_MOUSE	RK4_SPIOL	PSY_ARATH	FRA_DROME	CHRD_XENLA	MADI_HUMAN	Y144_MYCGE	VATA_METTH	ICAL_RABIT	RS7_MANSE	RS7_XENLA	PII1_MOUSE
-	-1	ч	-	-	٦	-	-	н	н	-	г
948	293	422	595	941	1425	279	584	718	190	194	332
30.1	29.8	29.8	29.8	29.8	29.8	29.5	29.5	29.5	29.1	29.1	29.1
ß	45	45	45	45	45	44.5	44.5	44.5	44	44	44
45.											

## ALIGNMENTS

g

```
REAUGULA AMINOPEDTIDASE

CAPSID PROTEIN C (POTENTIAL).

MAJOR ENVELOPE NOTEIN SI/E2 (POTENTIAL).

MAJOR ENVELOPE PROTEIN SI/E2 (POTENTIAL).

NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).

NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).

NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).

NONSTRUCTURAL PROTEIN NSI/E (POTENTIAL).

NONSTRUCTURAL PROTEIN NSI/E (POTENTIAL).

RA-DIRECTED RNA POLYMERASE (POTENTIAL).

RA-DIRECTED RNA POLYMERASE (POTENTIAL).

RA-DIRECTED RNA POLYMERASE (POTENTIAL).

ATP (POTENTIAL).

N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 146; DB 1; Length 3010;
Pred. No. 2.3e-13;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94A1C77435D642BB
                                                                  InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR00145; HCV_NS4a.
InterPro; IPR00145; HCV_NS4a.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR00166; HCV_RRP.
InterPro; IPR007094; RNA_Pol_DS_PS.
InterPro; IPR007094; RNA_Pol_DS_PS.
InterPro; IPR007094; RNA_Pol_DS_PS.
InterPro; IPR007094; HCV_COPE; I.
Pfam; PF01543; HCV_COPE; I.
Pfam; PF015549; HCV_NS1; I.
Pfam; PF01538; HCV_NS2; I.
Pfam; PF01006; HCV_NS3; I.
Pfam; PF01006; HCV_NS4b; I.
Pfam; PF01006; HCV_NS4b; I.
Pfam; PF01001; HCV_NS4b; I.
Pfam; PF01006; HCV_NS5a; I.
Pfam; PF01006; HCV_NS5a; I.
Pfam; PF000771; hellcase_C; I.
Pfam; PF000771; hellcase_C; I.
Pfam; PF000771; Hellcase_C; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           96.78;
96.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            326573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 96.7
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ¥,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      540
556
576
623
645
2041
2240
2529
3010 i
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3D-structure.
INIT_MET
                                            InterPro; InterP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement is (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-92230232; PubMed-1314459;
A Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
A Okamoto H., Fuxuai K., Okada S.I., Yamamoto K., Lizuka H.,
Tanaka T., Fuxuai K., Okada S.I., Yamamoto S.;
Tanaka T., Fuxuai S., Tsuda F., Mishiro S.;
Tanaka T., Fuxuai S., Tsuda F., Mishiro S.;
Tanaka T., Fuxuai S., Tsuda S., Tsuda S., Mishiro S.,
Tanaka T., Fuxuai S., Tsuda S., Mishiro S.,
Tanaka T., Fuxuai S., Suda S., Mishiro S., Nishiro S., Subuni T., The Virology INS A Rolle IN THE VIRAL RNA REPLICATION.
Subuni T. The Virologe. The Envelope Consists of Two Profeed By A Lipoprofein Envelope The Envelope Consists of Two Profeed By A Profeed Nishiro S., Nishiro 
                                                          01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2010 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (Ns) (Fragment).
Hepatitis C virus (isolate HC-37) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Þ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTEWIAL).
MATRIX PROTEIN (POTEWIAL).
MAJOR ENVELOPE PROTEIN E (POTEWIAL).
NONSTRUCTURAL PROTEIN NS1 (POTEWIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            · · · POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.

N.LINKED (GLCNAC.)

N.LINKED (GLCNAC.)
       513 AA
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002522; HCV.capsid.
InterPro; IPR002521; HCV.capsid.
InterPro; IPR002531; HCV.capsid.
InterPro; IPR002531; HCV.car.
InterPro; IPR002531; HCV.capsid.
InterPro; IPR002531; HCV.capsid.
IPRam; PP01543; HCV.capsid; I.
IPRam; PP01542; HCV.capsid; I.
IPRam; PP01542; HCV.capsid; I.
IPRam; PP01560; HCV.capsid; I.
IPRODA; PD186062; HCV.capsid; I.
IPRODA; PD186062; HCV.capsid; Cat professional Companion of Companion
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D10074; BAA00968.1; -.
   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11111;
POLG_HCVJ2
P27959;
                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
CHAIN
CHAIN
CHAIN
```

ö

Gaps

Gaps

ö

Indels

Length 520;

us-09-491-146a-23.rsp

SEQUENCE

တ္တ

셤

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CELLULAR ANINOPEPIDAEL
CAPSID PROTEIN C (POTENTIAL).
MATIX PROTEIN (CPOTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
BY SIMILARITY.
N-LINKED (GLCNAC. ...) (POTENTIAL).
                                                                                                                                                                                                                                                                     01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment).
Hepatitis C virus (isolate HCV-KF) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Abe K., Inchauspe G., Fullsawa K.;
"Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
-I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein; Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
INII_MET 1 1 REMOVED FROM CAPSID PROTEIN C
520
56499 MW; AA135246CF20D525 CRC64;
                                                        Score 144; DB 1;
Pred. No. 6.7e-14;
                                                                                                                                                                                                                                                      520 AA
                                                                                          0; Mismatches
                                                                                                                            28
                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. MEDLINE-93019030; PubMed-1383400;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR; JQ1925; JQ1925.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_caps1d; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01543; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D10687; BAA01529.1; -.
                                                    95.4%;
ilarity 96.4%;
Conservative
                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    115
1191
383
369
1196
233
305
418
 520 5
520 AA;
                                                    Query Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=31644;
                                                                                                                                                                                                                                                      POLG_HCVHK
Q01403;
 NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSMEM
                                                                                                                                                                                                                  RESULT 4
POLG_HCVHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
 FT
                                                                                                                            ö
                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMIS cutstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lab-sib.ch).
                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL)
                                                                       Gaps
                                                                                                                                                                                                                                                                 01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
GENVeLOPE 91ycoprotein El (GP32) (GP35); Envelope 91ycoprotein El (GP68) (GP69) (GP69)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus (isolate HCV-476) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE. CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                   Length 513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                       1; Indels
943F31E3514CDEF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC.)
                                   Score 144; DB 1;
Pred. No. 6.6e-14;
                                                                                                                                                                                                                                   520 AA.
                                                                   0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BY SIMILARITY
                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                            Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PP01542; HCV_core; 1.
Pfam; PP01542; HCV_core; 1.
Pfam; PF01549; HCV_core; 1.
55704 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D10688; BAA01530.1; -.
                                 Query Match 95.4%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                   STANDARD;
513 A.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1192
3384
3303
4418
4418
4418
4434
                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus
                                                                                                                                                                                                                               POLG_HCVH4
Q01404;
```

TRANSMEM CARBOHYD CARBOHYD

CARBOHYD CARBOHYD

INIT\_MET

CHAIN CHAIN CHAIN

CARBOHYD CARBOHYD

CARBOHYD CARBOHYD

K

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AI Protein.
REMOYED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/R2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/R2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
6enome polyprotein (Contains: Capsid protein C (Core protein); Matrix
Protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2) (Fragment).
Hepatitis C virus (isolate HC-J5) (HCV).
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                     Gaps
  N-LINKED (GLCNAC. .) (POTENTIAL)
N-LINKED (GLCNAC. .) (POTENTIAL)
N-LINKED (GLCNAC. .) (POTENTIAL)
                                                                                                                                                                                                                ;
                                                                                                                                                          Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein; Glycoprotein; Coat protein; Envelope protein; Transmembrane; Nonstructural protein.

INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C
                                                                                                                            Score 144; DB 1; Length 52.
Pred. No. 6.7e-14;
424 N-LINKED (GLCNAC. . .) (PC
431 N-LINKED (GLCNAC. . .) (PC
449 N-LINKED (GLCNAC. . .) (PC
520 5476 MW; 1D2BD0A6FF27349B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   737 AA
                                                                                                                                                                                                             0; Mismatches
                                                                                                                                                                                                                                                        1 PKPORKTKRNTIRRPODVKFPGGGOIVG 28
|||||||||||||||||||||||||||||||5
5 PKPQRKTKRNTNRRPQDVKFPGGGOIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL, D10075; BAA00969.1; -.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR00251; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PP01543; HCV_capsid; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01549; HCV_core; 1.
Pfam; PF01549; HCV_core; 1.
Pfam; PF01559; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                          95.4%;
96.4%;
                                                                                                                            Query Match
Best Local Similarity 95.4.
The conservative 
                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
424
431
449
520
520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             116
192
384
734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVJ5
                           CARBOHYD
CARBOHYD
NON_TER
SEQUENCE
  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                       RESULT 5
POLG_HCVJ5
ID POLG_H
AC P27960
DT 01-AUG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
  당달달당
                                                                                                                                                                                                                                                                                                            g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 24, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
Hepatitis C virus (isolate HC-J7) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ⋖,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VITOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL 3NA REPLICATION.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A MAD GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-92230232: PubMed-1314459; Okamoto K., Lizuka H., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                           (POTENTIAL).
                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 737;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 144; DB 1; Length 73
Pred. No. 9.8e-14;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                         81207 MW; 3AF699D82AD501B1 CRC64;
                                                                                                                                                                                               (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                 (GLCNAC.
                                                                                             (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                             737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLG_HCVJ7
P27961;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genotypes.
                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
NON_TER
SEQUENCE
                       CARBOHYD
                                                                                                                                         CARBOHYD
                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                              CARBOHYD
                                                                        CARBOHYD
                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVJ7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HIDDER TRANSPORTED DE CONTROL DE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22.-); Proteases/helicase NS3 (P70) (Heperivirin)
(EC 3.4.21.98); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P65); Nonstructural protein NS5A (P670) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                     Protein.
REWOVED FROM CADSID PROTEIN C BY THE
CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN ( POTEWIAL).
MATRIX PROTEIN (POTEWIAL).
MATRIX PROTEIN (POTEWIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTEWIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTEWIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; **Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-96235224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 144; DB 1; Length 737; Pred. No. 9.8e-14;
                                                                         protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                   N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE-91140698; PubMed-1847440;
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Probom; P0186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Co:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Virol. 65:1105-1113(1991).
                                                                                       Transmembrane; Nonstructural
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     95.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity 96.4%;
nes 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1487-1500.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       from human carriers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      649
737
737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVBK
                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                       TRANSMEM
                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                        INIT_MET
                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 7
POLG_HCVBK
ID POLG_H(
AC P26663,
DT 01-AUG
                                                                                                                                       CHAIN
                                                                                                                                                                       CHAIN
                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                      MEDLINE-98277846; PubMed-9568891;

X. MEDLINE-98277846; PubMed-9568891;

X. MEDLINE-98277846; PubMed-9568891;

A Yan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,

Steinkuehler C., Tomel, L., de Francesco R., Kuo L.C., Chen Z.;

"Complex of NS3 protease and NS4A peptide of BK strain hepatitis C

"Complex of NS3 protease and NS4A peptide of BK strain hepatitis C

"Virus: a 2.2-A resolution structure in a hexagonal crystal form.";

"Torial Sci. 7:837-847(1998).

"IL Protein Sci. 7:837-847(1998).

"IL PROPERION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE

HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

"CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6

position, Cys or Thi in Pl and Ser or Ala in Pl'.

"CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                            ď
                                                                                                                                                                                                                                                                                                                                                                                                    (RNA)(N):
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID SET OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                           X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.

WEDLINE-9701508B; Pubmed-8861916;

LOVE R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,

Moomaw E.W., Adachi T., Hostomska Z.;

"The crystal structure of hepatitis C virus NS3 proteinase reveals

trypsin-like fold and a structural zinc binding site.";

Cell 87:331-342(1996).
 phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pro; IPR007095; RNA_pol_DS_PS.
Pro; IPR007094; RNA_pol_PSvir.
PP01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1GX6; 10-APR-02.
1CUV; 26-JUN-00.
8OHM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               14-JAN-98.
08-APR-98.
15-NOV-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-NOV-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A38465; GNWVTC.
1A1Q; 25-MAR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF02907;
PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1031;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INS3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IGX5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 stam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PDB;
PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB)
```

HCV\_NS4a; ]

```
REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN NSI/EZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/EZ (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                             Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                          IAL)
                                                                                                                                                                                                                                            IAL)
                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                   POTENT
                                                                                                                                                                                                                         POTENT]
                                                                                                                                                                                                                                          POTENT
                                                                                                                                                                                                                               POTENT
                                                                                                                                                                                                                                     POTENT
                                                                                                                       POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                           N-LINKED GLCNAC.
                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                          N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                        LINKED
                                                                                                                                                                                                                                                       LINKED
                                                                                                                                                                                                                                                             LINKED
                                                                                                                                                                                                                                                                   LINKED
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SWART; SN00487; DEXDC; 1.
                                                                  645
2041
2041
2240
2240
2529
2788
2788
1035
1047
1050
                                                                                                                                                                                                                                                                                                                  076
                                                                                                                                                                                                                                                                                                                              085
092
094
097
                                                                                                                                                                                                                                                                                                                                                            1103
                                                                        3D-structure.
INIT_MET
                                                                                                                            ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                            CARBOHYD
                                                                                                                                                                       CARBOHYD
                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                  CARBOHYD
                                                                                                                        TRANSMEM
                                                                                                                                                                                                                                                                                                                  TURN
STRAND
HELIX
TURN
STRAND
TURN
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                        STRAND
STRAND
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                          HELIX
STRAND
                                                                                                                                                                                                                                                                                                      STRAND
                                                                                                                                                                                                                                                                                                                                                                                                      STRAND
                                                                                                                                                                                                                                                                                     STRAND
                                                                  CHAIN
                                                                        CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                   HELIX
TURN
TURN
                                                                                                                  CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P65); N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (RNA)(N).

SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-91088550; PubMed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S., Suyimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                Length 3010;
                                                                                                                      327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                              Score 144; DB 1; L
Pred. No. 4.7e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3010 AA
                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                       5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ.
                                                                                                                                                              Query Match 95.4
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
  1188
1197
1202
1204
1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISCUSSION OF SEQUENCE
1187 118
1189 119
1198 120
1203 120
1680 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                               HCVJA.
                                                                     TURN
STRAND
SEQUENCE
  TURN
STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                             RESULT B
POLG_HCVJA
    25558
                                                                                                                                                                                                                                                            ò
```

```
1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                       POLG_HCVH
P27958;
                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SMART; SW00487; DEXDC; 1.

W Polyprotein; diycoprotein; Transferase; RNA-directed RNA polymerase; M Polyprotein; Glycoprotein; Envelope protein; Helicase; ATP-binding; Core protein; Goat protein; Hydrolase, Serine protease.

T Transmembrane; Nonstructural protein; Hydrolase, Serine protease.

T TRANSMEM PROTEIN C POTENIN C BY THE CRELULIAR AMINOPEPTIDASE.

T CHAIN 115 CAPSID PROTEIN C (POTENIAL).

T CHAIN 1192 383 MAJOR ENVELOPE PROTEIN E (POTENIAL).

T CHAIN 384 729 NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).

T CHAIN 1007 1615 NON-STRUCTURAL PROTEIN NS1 (POTENTIAL).

T CHAIN 1616 1862 NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).

T CHAIN 1616 1862 NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).

T CHAIN 1863 2013 NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 144; DB 1; Length 3010;
Pred. No. 4.7e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC,
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
                                                                                                                                                                                                                                                                                                                    Interpro; IPR007095; RNA_pol_DS_ps.
Interpro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_caps.d; 1.
Pfam; PF01539; HCV_cenv; 1.
Pfam; PF01539; HCV_env; 1.
                                                                                                                          | IPR002519; HCV_env.
| IPR002531; HCV_NS1.
| IPR002518; HCV_NS2.
| IPR004109; HCV_NS3.
| IPR000745; HCV_NS4a.
| IPR001490; HCV_NS4a.
| IPR00168; HCV_NS5a.
| IPR002868; HCV_NS5a.
| IPR002166; HCV_RGRP.
                                                                                    HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF02907; HCV_NS3; 1.

PF01006; HCV_NS4s; 1.

PF01001; HCV_NS4b; 1.

PF01506; HCV_NS5s; 1.

PF00271; helicase_C; 1.

PF00998; Viral_RGRP; 1.

PF00998; Viral_RGRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             201.
2240
2788
327017 W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   95.48;
                                                              InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ¥
HSSP; P26663; 1JXP.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                             InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
```

ö

0; Mismatches

96.48;

Conservative

Best\_Local Similarity Matches 27; Conser

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions along as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
62-003 (Rel. 42, Last annotation update)
62-003 (Rel. 42, Last annotation update)
62-003 (Rel. 42, Last annotation update)
63-004 (GAPA) (GNI); Protein F1 (GPAB) (GPAB); Envelope glycoprotein E2 (GPAB) (GPAB) (GNI); Protein F7; Nonstructural protein NS2 (PPD) (Hepacivirin)
62-3-4-21-98); Nonstructural protein NSA (FAB); Nonstructural protein NSAB (PZ); Nonstructural protein NSAB (PZ); Nonstructural protein NSAB (PAC) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (Isolate H) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.; "Structure of the hepatitis C virus RNA helicase domain."; Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         oligonucleotide: the crystal structure provides insights into the mode of unwinding.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-9815421; Pubmed-9493370;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY PLIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
-1- PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
-1- SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39.
-1- SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 υ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Prince A.M.; "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; Pubmed-9187654;
STANDARD;
```

us-09-491-146a-23.rsp

```
Gaps
    . .) (POTENTIAL) . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3011 AA; 327142 MW; 772CBB29CCD94753 CRC64;
  N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 144; DB 1;
Pred. No. 4.7e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          95.4%;
nilarity 96.4%;
Conservative 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1636
1652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
27; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1584
1606
1614
1622
1627
1627
1635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLG_HCVJ6
ID POLG_HCVJ6
  CARBOHYD
CARBOHYD
STRAND
                                  TURN
HELIX
TURN
STRAND
TURN
HELIX
TURN
TURN
STRAND
HELIX
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
HELIX
TURN
STRAND
STRAND
HELIX
TURN
TURN
STRAND
STRAND
HELIX
TURN
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 10
   g
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                             Probom, PD186062; HCV_NS1; 1.
SWART; SM00487; DEXDc; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; APP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                              REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE.

CAPSID PROTEIN C.

ENVELOPE GIYCOPROTEIN E1.

ENVELOPE GIYCOPROTEIN E1.

PROTEIN P7.

NONSTRUCTURAL PROTEIN NS2.

PROTEASELFELICASE NS3.

NONSTRUCTURAL PROTEIN NS5A.

NONSTRUCTURAL PROTEIN NSTRUCTURAL PROTEIN NSTRUCTURAL PROTEIN NSTRUCTURAL PROTEIN NSTRUCTURAL PROTEIN NSTRUCTURAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINKED
LINKED
LINKED
                                                                                EMBL; M67463; AAA45534.1;
PIR; A36814; GNWVCH.
                                                    PDB; 1HEI; 25-NOV-98.
PDB; 1A1V; 16-FEB-99.
PDB; 1A1R; 17-JUN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

us-09-491-146a-23.rsp

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 1, Last annotation update)
03-FEB-2003 (Rel. 4), Last annotation update)
04-Ontains: Capsid protein C (Core protein) (P22);
05-050 (GP68) (GP10) (NS1); Protein P1; Nonstructural protein NS2 (P21)
05-050 (Rel. 41, 22, -1); Protease/helicase NS3 (P70) (Hepacivirin)
05-050 (Rel. 41, 22, -1); Protease/helicase NS3 (P70) (Hepacivirin)
05-050 (Rel. 41, 22, -1); Protease/helicase NS3 (P70) (Hepacivirin)
05-050 (Rel. 41, 22, -1); Protease/helicase NS3 (P70); Nonstructural protein NS54 (P65); Nonstructural protein NS58 (P66) (P70); Nonstructural protein NS58 (P66); Nonstructural protein NS58 (P60) (P70); Nonstructural protein NS58 (P60) (P70); Nonstructural protein NS58 (P60) (P70); Nonstructural protein NS58 (P60) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                            CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELORE PROTEIN (POTENTIAL).
MAJOR ENVELORE PROTEIN (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTESAE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                     Proboni; relevee;
SMART; SMO4087; DEXDC; 1.
SMART; SMO4087; DEXDC; 1.
SCAPE PO1yprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Goat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE
CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                  SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ( POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MW; F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                    (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. 4.7e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 144; DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                 ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        95.4%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus.
NCBI_TaxID-11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLG_HCVJ8
P26661;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                 CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 11
POLG_HCVJ8
                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities regulres a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                         Genome polygrotein [Contains: Capsid protein C (Core protein) (P22); Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP32) (GP35); Envelope glycoprotein E2 (GP38) (GP70) (NS1); Protein P7; Nonstructural protein NS24 (P42); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Nonstructural protein Hepatitis C virus (isolate HC-J5) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
            01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002868; HCV_NS5a.
InterPro; IPR002166; HCV_EARP.
InterPro; IPR0010560; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
Pfam; PF011643; HCV_Capsid; 1.
Pfam; PF011642; HCV_Core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002531; HCV_NS1.
InterPro; IPR004109; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR000745; HCV_NS4a.
InterPro; IPR001499; HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D00944; BAA00792.1; -. PIR; JQ1303; JQ1303. HSSP; P27959; 1HEI. MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002519;
                                                                                                                                                                                                                               NCBI_TaxID=11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01006;
PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF02907;
```

ö

Gaps

Pfam; Pfam; Pfam; Pfam;

Pfam; Pfam;

```
MEDLINE-92230206; PubMed-1314449;
MEDLINE-92230206; PubMed-1314449;
MEDLINE-92230206; PubMed-1314449;
MEDLINE-92230206; PubMed-1314449;
MEDLINE-92230206; PubMed-1314449;
MEDLINE-9223026; Lin R.F., Lin P.C., Lin C.J., Chen D.S.;
Medline P.J., Lin M.H., Tai M.F., Lin P.C., Lin C.J., Chen D.S.;
Mapping the 5' termin of viral genomic and antigenomic RNA.";
Virology 188:102-13(1992). It genomic and antigenomic RNA.";
Virology 188:102-13(1992).
MEDLINE SMALL PROTEINS NS2A, NS2A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSTEBLE MEMBRANE-RELATED FUNCTION.
MEDLINE RELATED FUNCTION.
CONTAINTY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CONTAINTY AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA RNA PART R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Last annotation update)
16-March 12, Last annotation 18-March 18
                             RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                               SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 144; DB 1; Length 30 Pred. No. 4.7e-13; 0; Mismatches 1; Indels
                                                                                                              (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MW; 1A173E7E3381FD1A
                                                                                                     CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLCNAC
                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLCNAC
                                                                                                                                                                                                                                ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                N.LINKED
N.L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ch 95.4%;
1 Similarity 96.4%;
27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2038
2359
2811
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 12
POLG_HCVTW
ID POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                 ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                   TRANSMEM
                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P29846;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      between the Swiss Institute of Bioinformatics and the EMBL outstation—
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
W Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding;
W Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
T INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
T CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).
T CHAIN 192 383 MAJOR ENVELOPE PROTEIN B. (POTENTIAL).
T CHAIN 734 1010 NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
T CHAIN 1610 NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
T CHAIN 1619 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1620 1866 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1620 1866 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1620 1866 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
                                                                                                                                                                                                                                    SEQUENCE FROM N.A. MEDILINE-92230212; PubMed-1314459; Okamoto H., Kural K., Okada S. I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGNS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Interpro; IPR002868; HCV_NSSa.
Interpro; IPR007195; HCV_RGRP.
Interpro; IPR007095; RNA_DO1_DS_PS.
Interpro; IPR007094; RNA_DO1_PSvir.
Pfam; PF01543; HCV_caps1d; 1.
Pfam; PF01543; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pr0Dom; PD166062; HCV_NS1; 1.
Pr0Dom; PD166062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJB. HSSP; P27598; 1HEI. MRROPS; S29.001; -. MEROPS; 039.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR004109; H
IPR000745; F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF01539; Pfam; PF01560; Pfam; PF01538; Pfam; Pf01538; Pf0153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ÷
```

ö

Gaps

```
EMBL; M62321; AAA45676.1; -.
                                                                                                                                       Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11104;
                                                                 645
2041
2077
2240
2529
2788
3010
                                                                                                                                                                                                                                                   POLG_HCV1
P26664;
                                     CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                     CARBOHYD
CARBOHYD
SEQUENCE
           CARBOHYD
CARBOHYD
                                                                                             CARBOHYD
                                                                                                                                                                                                                                 RESULT 13
                                                                                                                                                                                                                                           POLG_HCV1
 g
                                                                                                                                                                                                                                                    à
                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CELLULAR AMINOPEPTIDASE.

CORE PROTEIN (POTENTIAL).
MAJOR ENVELOR (POTENTIAL).
NONSTRUCTURAL PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLXMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Fransmembrane; Nonstructural protein; Hydrolase; Serine protease;
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.

SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
                                                                                                               EMBL; M84754; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                     InterPro; IPR002166; HCV_RdRP.
InterPro; IPR007094; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_PSvir.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                 HCV_capsid.
HCV_core.
                                                                                                                                                                                                Interpro; IPR002519; HCV_NS1.
Interpro; IPR002518; HCV_NS1.
Interpro; IPR004109; HCV_NS2.
Interpro; IPR000745; HCV_NS3.
Interpro: IPR000745; HCV_NS4a.
                                                                                                                                                                                                                                         InterPro; IPR000745; HCV_NS4a.
InterPro; IPR001490; HCV_NS4b.
InterPro; IPR002868; HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                         ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                     PF01539; HCV_env; 1.
PF01560; HCV_NS1; 1.
PF01538; HCV_NS2; 1.
PF02907; HCV_NS3; 1.
PF01006; HCV_NS4a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1006
1615
1862
22013
3010
1083
11103
11165
1237
1237
1319
209
2234
2254
417
423
                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00487; DEXDC;
                                                                                                                                 PDB; 1N64; 25-FEB-03.
PDB; 1NS3; 08-APR-98.
                                                                                                                                                                                InterPro; IPR002522;
                                                                                                                                                                       InterPro; IPR001410;
                                                                                                                                                                                            InterPro; IPR002521;
                                                                                                                       PIR; A40244; GNWVTW
                                                                                                                                                    MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF01542; PF01539; PF01560; P
                                                                                                                                                                                                                                                                                                                                                                                                                                                         3D-structure
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSMEM
                                                                                                                                                                                                                                                                                                                                                           Pfam; Pfam; Pfam; P
                                                                                                                                                                                                                                                                                                                                Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics in Stitute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP69) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.); Proteases/Helicase NS3 (P70) (Hepaclylin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate 1) (HCV).
Viruses; SBNNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (RNA)(N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                (POTENTIAL)
    (POTENTIAL)
                                                                             POTENTIAL)
                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC. .) (POTENTIAL)
MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 137; DB 1;
Pred. No. 5.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                            (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                             (GLCNAC
                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3011 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 KPORKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                     623
645
645
1 2041
2077
2240
2529
2788
1, 327047 M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          90.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE-88288341; PubMed-2840577;
Perrine K.G., Denker J.A., Nilsen T.W.;
"A multi-copy gene encodes a potentially protective antigen in Brugia
                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nilsen T.W., Maroney P.A., Goodwin R.G., Perrine K.G., Denker J.A., Nadduri J., Kazura J.W., "Cloning and characterization of a potentially protective antigen in Proc. Natl. Acad. Sci. U.S.A. 85:3604-3607(1988).
                                                                                                                                                                                                    PSTN_BRUMA STANDARD; PRT; 548 AA.
P10723;
01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
15-SEP-2003 (Rel. 42, Last annotation update)
Asparaginyl-tRNA synthetase, cytoplasmic (EC 6.1.1.22) (Asparagine-tRNA igase) (AsnRS) (Potentially protective 63 kDa antigen).
Brugia malayi (Filarial nematode worm).
Bukaryota; Mctazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
NCBI_TAXID-6279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-96049569; Pubmed-7589498;
Kron M., Marquard K., Hartlein M., Price S., Leberman R.;
"An immunodominant antigen of Brugia malayi is an asparaginyl-tRNA
                                                               ö
                               Length 3011;
327197 MW; 65F8C9447FCE5AF9 CRC64;
                               Score 136; DB 1;
Pred. No. 7.4e-12;
                                                              1; Mismatches
                                                                                            1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                              Mol. Biochem. Parasitol. 30:97-104(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro: IPR004522; AsnS.
InterPro: IPR004522; RNA-synt_2
InterPro: IPR004364; tRNA-synt_asp.
InterPro: IPR004365; tRNA_ant1.
InterPro: IPR006195; tRNA_lgase_II
Pfam; PF00152; tRNA_synt_2; I
Pfam; PF01365; tRNA_shr1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; J03971; AAA27852.1; --
EMBL; J03266; AAA27849.1; --
PIR; A28209; A28209.
PIR; A54510; A54510.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
MEDLINE-88217950; PubMed-3368467;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TIGR00457; asnS; 1.
                            Query Match 90.1%;
Best Local Similarity 89.3%;
Matches 25; Conservative
3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FIGREAMS;
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FUNCTION.
                                                                                                                                                                         RESULT 14
SYN_BRUMA
g
                                                                                             ò
                                                                                                                        g
                                                                                                                                                                                                                         CELLUIAR AMINOPEPTIDASE.
CAPSID PROTEIN (POTENTIAL).
MAJRIX PROTEIN (POTENTIAL).
NONSTRUCTURAL PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE RELAY SYSTEM (BY S. CHARGE RELAY SYSTEM (BY S. CHARGE RELAY SYSTEM (BY S. ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                InterPro; IRPR002519; HCV_env.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR004109; HCV_NS2.
InterPro; IPR004109; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS5A.
InterPro; IPR001650; HCV_NS5A.
InterPro; IPR001650; HCV_RARP.
InterPro; IPR007095; RNA_POL_DS_PS
InterPro; IPR007095; RNA_POL_DS_PS
InterPro; IPR007095; RNA_POL_PSvir.
Pfam; PF01543; HCV_caps1d; 1.
Pfam; PF01543; HCV_caps1d; 1.
Pfam; PF01550; HCV_NS1; 1.
Pfam; PF01550; HCV_NS1; 1.
Pfam; PF01006; HCV_NS2; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01050; HCV_NS5; 1.
Pfam; PF01050; HCV_NS4; 1.
Pfam; PF01050; HCV_NS5; 1.
                                                                    InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00998; Viral_R(ProDom; PD186062; HCV_1
SMART; SM00487; DEXDC;
          1A1V; 16-FEB-99
1HEI; 25-NOV-98
                                        MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       347
1083
11165
11165
11230
1230
209
234
305
417
423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    and genetic sequences encoding a receptor-like protein.";
EMBO J. 5:2197-2202(1986).
PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Leduc R., Molloy S.S., Thorne B.A., Thomas G.;
"Activation of human furin precursor processing endoprotease occurs
an intramolecular autoproteolytic cleavage.";
J. Biol. Chem. 267:14304-14308(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              strake A.J.;
cons and gene structure for a human subtilisin-like protease with
cleavage specificity for paired basic amino acid residues.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FURI_HUMAN STANDARD, PRT; 794 AA.
P09958, Q14336;
01-MAPL1989 (Rel. 10, Created)
01-MAPL1990 (Rel. 14, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
15-SEP-2003 (Rel. 42, Last annotation acid residue cleaving enzyme) (PACE) (Dibasic processing enzyme).
FURIN OR PACE OR FUR.
HOMO sapiens (Human).
                                                                                       IMMUNOLOGICAL REACTIVITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      van den Ouweland A.M.W., van Duijnhoven H.L.P., Keizer G.D., Dorssers L.C.J., van de Ven W.J.M.;
"Structural homology between the human fur gene product and the subtilisin-like protease encoded by yeast KEX2.";
Nucleic Acids Res. 18:664-664(1990).
                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Roebroek A.J.M., Schalken J.A., Leunissen J.A.M., Onnekink C., Bloemers H.P.J., van de Ven W.J.M.;
Evolutionary conserved close linkage of the c-fes/fps proto-o
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Barr P.J., Mason O.B., Landsberg K.E., Wong P.A., Kiefer M.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Slezen R.J., Creemers J.W.M., van de Ven W.J.M., lendlogy modelling of the catalytic domain of human furin. for the eukaryotic subtilisin-like proprotein convertases."; Eur. J. Blochem. 222:255-266(1994).
                                                                                                                                                                                                                                                                                                              DB 1; Length 548;
                                                                                                                                                                                                                                                                                                                                                                      10; Indels
                                                                              REGION OF IMMUNOLOGICAL REAK

1 -> V (IN REF. 2).

P -> S (IN REF. 2).

L -> V (IN REF. 2).

L -> V (IN REF. 2).

C -> K (IN REF. 2).

W; 7E4F6EA248255095 CRC64;
                                                                                                                                                                                                                                                                                                                                                                   4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -STRUCTURE MODELING OF CATALYTIC DOMAIN.
                                                                                                                                                                                                                                                                                                           Score 54;
Pred, No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORDAODNTLTESVDLLMPGVGEIVG 474
                                                                                                                                                                                                                                                                                                                                                                                                                          4 QRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                        Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 296-794 FROM N.A. MEDLINE*87053858; PubMed=3023061;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
MEDLINE-91321735; PubMed-1713771;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-92332543; PubMed-162922;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-90175002; PubMed-2408021;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        cleavage specificity for paired DNA Cell Biol. 10:319-328(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=94291619; PubMed=8020465
                                                                                 131 RE
50 I
65 H
212 P
265 L
454 O
62339 MW;
                                                                                                                                                                                                                                                                                                        sh 35.8%;
| Similarity 44.0%;
| 11; Conservative
                                                                              81
50
65
212
265
265
454
548 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE-Blood
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROCESSING
                                                                                                                                                                CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              450
                                                                                                                                                                                                                                                                                                              Query Match
                                                       Antigen.
                                                                                                                                                                                                                                                                                                                                        Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Brake A
                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ULT
  KW
KW
KW
FT
FT
FT
FT
SQ
                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       THE SET OF THE SET OF
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
FUNCTION: FURIN IS LIKELY TO REPRESENT THE UBIQUITOUS ENDOPROTEASE CACTUATY WITHIN CONSTITUTIVE SECRETORY PARHWAYS AND CAPABLE OF CLEAVAGE AT THE RX(K/K)R CONSENSUS MOTIF.
CATALYTIC ACTIVITY: Release of mature proteins from their proproteins by cleavage of Arg-Yaa-Yaa-Arg-1-Zaa bonds, where Xaa can be any amino acid and Yaa is Arg or Lys. Releases albumin, complement component C3 and von Willebrand factor from their
                                                                                                  COFACTOR: CALCIUM-DEPENDENT.
ENZYME REGULATION: COULD BE INHIBITED BY THE NOT SECONDLY CLEAVED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  protease; Transmembrane; Glycoprotein; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MIM; 136930; ...
GO; GO:0005794; C:Golgi apparatus; TAS.
GO; GO:0004276; F:furin activity; TAS.
GO; GO:0007267; P:cell-cell signaling; TAS.
GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHARGE RELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CYS-RICH.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF01483; P_proprotein; PARTIAL. Pfam; PF00082; Peptidase_S8; 1. PRINTS; PR00723; SUBTILISIN. ProDom; PD000717; P_domain; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROSITE; PS00136; SUBTILASE_ASP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUBTILASE_HIS; 1.
SUBTILASE_SER; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FURIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR006212; Furin_repeat.
InterPro; IPR002884; P_domain.
InterPro; IPR000209; Peptidase_S8.
                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; X17094; CAA34948.1; -. EMBL; X04329; CAA27860.1; -. EMBL; A06939; CAA00605.1; -. PIR; A39552; KXHUF.
                                                                                       respective precursors.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genew; HGNC:8568; FURIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FU;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1MPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROSITE; PS00137;
PROSITE; PS00138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hydrolase; Serine
Zymogen; Calcium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEROPS; S08.071;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00261;
                                                                                                                           PROPEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         099405;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              136950
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IGNAL
```

```
ö
                                                                                                                      Gaps
                                                                                                                      ;
0
70 75 CLEAVAGE (SECOND AUTO-).
104 107 CLEAVAGE (FIRST AUTO-).
498 500 CELL ATPACHMENT SITE (POTENTIAL).
759 762 CELL SURFACE SIGNAL.
773 779 TRANS GOLGI NETWORK SIGNAL.
774 AA; 86678 MW; 10C44DD5892EF85D CRC64;
                                                                                         Score 52; DB 1; Length 794;
Pred. No. 6.8;
4; Mismatches 5; Indels
                                                                                         Query Match
Best Local Similarity 50.0%;
Matches 9; Conservative
 SITE
SITE
SITE
SITE
SITE
 FF FF S
```

4 QRKTKRNTIRRPQDVKFP 21 :|:|||: | | ||| 102 KRRTKRDVYQEPTDPKFP 119 g

ò

Search completed: August 7, 2003, 11:19:55 Job time : 6.90909 secs

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

0.08bcx1 0.08bcx1 0.08bcx0 0.08bcw7 0.08bcw7 0.08drJ1 0.0

hepatitis

hepatitis hepatitis

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

hepatitis

Q8qm14 Q68892 Q68893

Perfect score:

Sedneuce:

protein

₹

Run on:

Scoring table:

Searched:

Minimum DB s Maximum DB s

Database :

```
VITUB.;
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
I-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN MAND GIXCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN MAND GIXCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
EMBL, U23745; AAA56052.i...
InterPro; IPR002522; HCV_caps1d.
Pfam, PF01543; HCV_caps1d.
                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                            Songsivilai S., Kanistanon D., Kunkitti R.; "Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   45
5014 MW; CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                        Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence upd:
01-OCT-2002 (TrEMBLrel. 22, Last annotation upd:
01-OCT-2002 (TrEMBLrel. 22, Last annotation upd:
01-OCT-2004 (TremBlrel. 22, Last annotation upd:
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                45 AA.
                                                                                                                                                                                                                                                                      ALIGNMENTS
                                 Q81340
Q8QMJ4
Q68892
                                                                                                                              08BCW4
08QRJ1
08QMI6
                                                                                                                                                                        Q8QRJ5
Q8QMI5
Q8QMI9
Q8QMJ1
                                                                                                                                                                                                                            Q8QRJ3
Q8QRI8
Q8QRI9
                                                                           Q8BCX1
Q8BCX0
                                                                                             QBBCW9
QBBCW8
                                                                                                             QBBCW7
QBBCW6
                                                                                                                                                       OBOMIB
OBORJ4
                                                                                                                                                                                                           Q8QMJ0
Q8QRI7
                                                           Q68893
Q8BCX2
                          080P86
                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. STRAIN-HCV-BB21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           45 AA;
Hepacivirus,
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein.
SEQUENCE
                                                                                                                                                                                                                                                                                                                068307
                                                                                                                                                                                                                                                                                               RESULT 1
Q68307
hepatitis hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          hepatitis
                                                        August 7, 2003, 11:05:41; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                  is the number of results predicted by chance to have a neer than or equal to the score of the result being printed,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       066308
068310
068310
0819x32
0819x4
0840P32
0840P72
0840P77
0840P77
0840P77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               068307
                                                                                                                                                                                                                                                                                                                                                                                                                            score greater than or equal to the score of the result bein and is derived by analysis of the total score distribution.
       GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                       al number of hits satisfying chosen parameters:
                                                                                                    151
1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                      830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q68310
Q8JYS2
Q8JYS3
Q8JYR4
Q8QP85
                                                                                                                                                                                                                                                                                                                                                                 sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                      sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                              BLOSUM62
Gapop 10.0 , Gapext
                                                                                                                                                                                                                                                                                                                                                                                  sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                     sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                           sp_archea:*
sp_bacteria:*
sp_fungi:*
                                                                                                                                                                                         seq length: 0
seq length: 200000000
                                                                                            US-09-491-146A-23
                                                                                                                                                                                                                                                                                                                                                                                                    sp_archeap:*
                                                                                                                                                                                                                                                                                                                                       sp_plant:*
sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                222222222222222
                                                                                                                                                                                                                                                                                                                                                          sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query
Match Length DB
                                                                                                                                                                                                                                                    SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100.0
```

ö

Gaps

ö

Indels

; 0

Mismatches

ilarity 100.0%; P. Conservative 0;

Best Local Similarity Matches 28; Conser

hepatitis hepatitis

080P72 080P71 080P74

100.0 100.0 100.0

100.0

Pred. No.

Result ŝ

080P84 080P75 080P79 080P80 080P77

100.0

hepatitis hepatitis

Query Match

1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28 

g à

> hepatitis hepatitis hepatitis hepatitis

100.0%; Score 151; DB 12; Length 45; 100.0%; Pred. No. 1e-15;

K

RESULT 2 Q68308

```
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                        Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
EMBL, AF506613; AAM33389-1; -
InterPro; IPR002522; HCV_capsid.
Pfam: PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Generic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/Genenk-DDB Jatebases.
-!- SUBUNIT: THE VIRION OF THIS YIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN MAND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN MAND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
EMBL, AF506612; AAM33381.; -
InterPro: IPR002522; HCV_capsid.
Pfam: PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100.0%; Score 151; DB 12; Length 60; 100.0%; Pred. No. 1.4e-15; Live 0; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         60 AA; 6725 MW; ACCE7D9C8B90299A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      B92CCE7D9C8B9029 CRC64;
                                                                                                                                                               (TrEMBLrel. 22, Created)
(TrEMBLrel. 22, Last sequence update)
(TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 AA.
                                                                                                                                  60 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CORE PROTEIN
1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                   1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-OCT-2002 (TrEMBLrel. 22, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                            Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
6881 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-OCT-2002 (TrEMBLrel. 22, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                         01-MAR-2003 (TrEMBLrel, 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 100.
nes 28; Conservative
                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A. STRAIN-RIG256;
                                                                                                                                                                                                                                                                                Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                          STRAIN-RIG253;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein.
CHAIN
                                                                                                                                                                 01-0CT-2002
01-0CT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                08JYS2
08JYS2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8JYS3;
Q8JYS3;
                                                                                          RESULT 4
Q8JYS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QBJYS3
                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                            Songsivilai S., Kanistanon D., Kunkitti R.; "Identification and characterisation of Thai isolates of hepatitis C
                                                                                                                                                                                                                                                                                                                                                          K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Songsivilai S., Kanistanon D., Kunkitti R., Isolates of hepatitis {\tt C} "Identification and characterisation of Thai isolates of hepatitis {\tt C}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MINA (BY SIMILARITY).

EMBL; U23748; AAA65055.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                             Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
--- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
--- PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF EMBL; U23746; AAA65053.1;
--- InterPro; IPR002522; HCV_capsid.
--- Pfam; PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
01-OCT-2011 (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 12; Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100.0%; Score 151; DB 12; Length 45; 100.0%; Pred. No. 1e-15; Live 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein.
NON_TER 45 45
SEQUENCE 45 AA; 5014 MW; CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        45 45 45
45 AA; 5014 MW; CC527167096DDAF6 CRC64;
                                                   ol-NoV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0%; Score 151; DB 12;
100.0%; Pred. No. 1e-15;
tive 0; Mismatches 0;
                                   45 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPORKTKRNTIRRPODVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1996 (TrEMBLrel. 01, Created)
                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
es 28; Conserv
                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-HCV-C94009;
                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                              STRAIN-HCV-BB38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein.
                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       068310
                             908890
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             virus.
```

**Matches** 

DR DR DR

RESULT 3 068310

OPPTE

셤

ð

ö

Gaps

ö

Ø

K

Gaps

```
STRAIN-233;
MEDLINE-2194745; PubMed-11907242;
Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
A natural intergenotypic recombinant of hepatitis C virus identified
in St. Petersburg.";
J. Virol. 76:4634-4643(2002).
I. VIROL. 76:4634-4643(2002).
PROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; AYO70178; AAL58592.1;
InterPro: IPR002522; HCV_capsid.
Prant. PP01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
"A natural intergenotypic recombinant of hepatitis C virus identified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                              ö
                                                                                               Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                100 100
100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100
11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JUN-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last annotation update)
                                                                            Score 151; DB 12;
Pred. No. 2.5e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0%; Score 151; DB 12;
ilarity 100.0%; Pred. No. 2.5e-15;
Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Core protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                            100 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100 AA.
                                                                                           Match
Local Similarity 100.0%; Score 151; I
Local Similarity 100.0%; Pred. No. 2.5
les 28; Conservative 0; Mismatches
                                                                                                                                                           1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                               01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last seq
01-OCT-2002 (TrEMBLrel. 22, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-609;
MEDLINE-21904745; Pubmed-11907242;
InterPro; IPR002522; HCV_caps1d.
Pfam; PF01543; HCV_caps1d; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                            Polyprotein.
NON_TER 100
SEQUENCE 100 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
NON_TER 10
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          080P72
                                                                                                                                                                                                                                                                          Q8QP87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           080P72
                                                                                                                 Best Loca
Matches
                                                                                                                                                                                                                                           æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 9
                                                                                                                                                                                                                                         RESULT 6
Q8QP87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q8QP72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DA
ET
SO
FT
SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
"A natural intergenotypic recombinant of hepatitis C virus identified
in St. Petersburg.";
-1. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN BNVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; AY070180; AAL58594.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                   Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DBBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROFIIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL, AF506621; ARAM33997.1; -
InterPro; IPR002522; HCV_capsid.
Prim: PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                         Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                               ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0%; Score 151; DB 12; Length 73; 100.0%; Pred. No. 1.8e-15; ive 0; Mismatches 0; Indels (
 Length 61;
                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6EDC082DAE8CB796 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Last sequence update)
Last annotation update)
                                                                                                                                                                                                                            01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
100.0%; Score 151; DB 12;
100.0%; Pred. No. 1.4e-15;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-JUN-2002 (TrEMBLrel. 21, Last sequence upda
01-027-2002 (TrEMBLrel. 22, Last annotation up
Core protein (Genome polyprotein) (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100 AA.
                                                                                                                                                                             73 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                01-OCT-2002 (TrEMBLrel. 22, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-602;
MEDLINE-21904745; PubMed-11907242;
                                                                                                                                                                                                                                                          Genome polyprotein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 73 73
73 AA; 8201 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-JUN-2002 (TrEMBLrel. 21,
01-JUN-2002 (TrEMBLrel. 21,
01-OCT-2002 (TrEMBLrel. 22,
Ouery Match
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity 100.
Les 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                      STRAIN-RIG291;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
                                                                                                                                                                                                                          01-OCT-2002
                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                           QBJYR4
QBJYR4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q8QP85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                          RESULT 6
QBJYR4
                                                                                             g
                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                              100 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100 AA;
                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11103;
         NCBI_TaxID=11103;
                                                                                                                                                                                                                  100
                                                                                                                                                                                                        Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                         Q8QP80;
                                                                                                                                                                                                                                                                                                                                                                                            089090
                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 13
                                                                                                                                                                                                                                                                                                                                                                               08QP80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 080P77
         δ
                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-21904745; PubMed-11907242;
Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
A natural intergencypic recombinant of hepatitis C virus identified
in St. Petersburg ";
J. Virol. 76:4034-4043(2002).
-!- SUBMINIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
--- IIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA BY SIMILARITY).
EMBL, ANO70194, AAL58608.1;
----
INTERPOOR TO THE VIRUS IS SIMILARITY).
FEMEL, ANO70194, AAL58608.1;
----
INTERPOOR TO THE VIRUS IS SIMILARITY).
in St. Petersburg.";
J. Virol. 76:4034-4043(2002).
I. Virol. 76:4034-4043(2002).
I. SUBUNIT. THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN THE VIRION OF THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL, AYO70193; AALSB607.1; -
InterPro; IPR002522; HCV_capsid.
Pfam. PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                               100.0%; Score 151; DB 12; Length 100; 100.0%; Pred. No. 2.5e-15; 1.ve 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLYPIOLEIN.
NON_TER 100 100
SEQUENCE 100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                      SEQUENCE 100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                          01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JUN-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JUN-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; Score 151; DB 12;
; Pred. No. 2.5e-15;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                     100 AA
                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                 100.0%; Pre-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 100.0%;
Matches 28; Conservative (
                                                                                                                                                                                       28; Conservative
                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                             Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
                                                                                                                            100
                                                                                                            Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-637
                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                   Q8QP71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q8QP74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11

080P74

ID 080P7

AC 080P7

DT 01-JU

DT 01-JU

DT 01-OC

DE CORE

OS HEPAL
                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                           SULT 10
   ò
                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                    S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
C STRAIN-631;

X MEDLINE-21904745; PubMed-11907242;

X Ralinina O., Norder H., Mukomolov S., Magnius L.O.;

Ra Andrural intergency procession of hepatitis C virus identified in St. Petersburg.";

Thist. Petersburg.";

YICOL. 76:4034-4043(2002).

CC ILPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA BY SIMILARITY).

CC PROTEIN C AND MRNA BY SIMILARITY).

DR EMBL; AYO'00191; AALS8605.1; ---

DR FMBL; AYO'0191; AALS8605.1; ---

DR FMBL; AFO'0191; AALS8605.1; ---

DR FMBL; PFO'143; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-21904745; PubMed-11907242;
Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
A natural intergenctypic recombinant of hepatitis C virus identified
in St. Petersburg...;
J. Virol. 76:4034-4043(2002)...
I. POBDNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
ILPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VIruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0%; Score 151; DB 12; Length 100; 100.0%; Pred. No. 2.5e-15; live 0; Mismatches 0; Indels 0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0%; Score 151; DB 12; Length.100; 100.0%; Pred. No. 2.5e-15; 1.ve 0; Mismatches 0; Indels 0.
                                                                                                                                                                                                                                                                                                                                                                                                                                         11289 MW; F3BC90F5F595F18B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100
11289 MW; F3BC90F5F595F18B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence upds
01-OT-2002 (TrEMBLrel. 22, Last annotation up
01-OCT-protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPORKTKRNTIRRPODVKFPGGGOIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; AY070185; AAL58599.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q8QP77
Q8QP77;
```

us-09-491-146a-23.rspt

```
Search completed: August 7, 2003, 11:18:55 Job time: 28.6364 secs
                                                                                                                                                                                                                                                                               Query Match
                                         080P75
                         RESULT 15
                                                 ç,
                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                               Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100.0%; Score 151; DB 12; Length 100; 100.0%; Pred. No. 2.5e-15; 1ve 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                       100.0%; Score 151; DB 12; Length 100; 100.0%; Pred. No. 2.5e-15; tive 0; Mismatches 0; Indels 0
                                                                                                                                                                                                 100 100
100 AA; 11289 MW; F3BC90F5F595F18B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-CT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-CCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                100 AA
                                                                                                                                                                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                               Best Local Similarity
Matches 28; Conserv
                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
NON_TER 1
SEQUENCE 10
                                                                                                                                                                                        Polyprotein.
                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                 Q8QP84
                                                                                                                                                                                                                                                                                                RESULT 14
Q8QP84
                                                                                                                                                                                                                                                                                                                  CSSSEPPFS
                                                                                                                                                                                                                                                                                                                                                                                                                STED
                                                                                                                                                                                                                                                         ð
```

ö

Gaps

0

1 PKPORKTKRNTIRRPODVKFPGGGGIVG 28

ò

```
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                    01-JUN-2002 (TrEMBLrel. 21, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100.0%; Score 151; DB 12;
100.0%; Pred. No. 2.5e-15;
tive 0; Mismatches 0;
                                                                                                                                                    Core protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                      01-JUN-2002 (TrEMBLrel. 21, Created)
                                                                                                                                                                                                                                                          STRAIN-680;
MEDLINE-21904745; Pubmed-11907242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 100.
Matches 28; Conservative
                                                                          PRELIMINARY;
                                                                                                                                                                                                                                [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                           NCBI_TaxID-11103;
```

```
Query Match 100.0%; Score 151; DB 3; Best Local Similarity 100.0%; Pred. No. 3.3e-15; Matches 28; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOLECULE TYPE: protein HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Her
US-08-921-887-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                              US-08-921-887-23
   22666666666666444444
86616666666664444444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 23, Appl
Sequence 187, App
Sequence 188, App
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Appli
, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Appli
Appli
Appli
Appli
Appli
                                                                                                          August 7, 2003, 11:07:41; Search time 10.5455 Seconds (without alignments) 112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 25, Sequence 191, Sequence 175, Sequence 175, Sequence 175, Sequence 6, A Sequence 2, A Sequence 2, A Sequence 2, A Sequence 2, S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence Sequence Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Issued_Patents_AA:*
.: /cgn2_6/ptodata/1/laa/5A_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/5B_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/6A_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/6B_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/PcTUS_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/PcTUS_COMB.pep:*
             GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-921-887-23
US-08-290-665A-187
US-08-290-665A-188
US-08-290-665A-198
PCT-US95-10398-187
PCT-US95-10398-198
PCT-US95-10398-199
US-08-635-886C-230
US-08-635-886C-192
US-08-635-886C-192
US-08-635-886C-192
US-08-635-886C-192
US-08-635-886C-192
US-08-635-886C-192
US-08-635-886C-192
US-08-635-10398-175
                                                                                                                                                                                                                                                                                                                             al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PCT-US93-08638-9
PCT-US94-07088-23
                                                                                                                                                                           US-09-491-146A-23
151
1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                              protein search, using sw model
                                                                                                                                                                                                                                              BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                 length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100
                                                                                                                                                                                                                                                                                                                                                                 sed
                                                                                                                                                                                            Perfect score:
                                                                                                                                                                                                                                              Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score
                                                                                OM protein
                                                                                                                                                                                                                                                                                                                                                               Minimum DB
Maximum DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Database :
                                                                                                                                                                                                               Sequence:
                                                                                                                                                                                                                                                                                                Searched:
                                                                                                                 Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
```

```
232, App
14, Appl
8, Appli
8, Appli
8, Appli
8, Appli
2, Appli
15, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Appl
Appl
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                    2, 1
15, 4
46,
                                                                Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence Sequence Sequence Sequence Sequence 1
US-08-835-075A-10
US-08-635-886C-198
US-08-835-075A-14
US-08-34-977-8
US-08-34-977-8
US-08-34-616-8
US-09-315-850-8
US-08-311-850-8
US-08-851-195-2
US-08-851-195-1
US-08-851-195-1
US-08-851-195-1
US-08-851-195-1
US-08-851-195-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PCT-US95-13552-2
US-07-681-701-16
US-08-078-271B-1
```

## ALIGNMENTS

```
Sequence 23, Application US/08921887
Sequence 23, Application US/08921887
Patent No. 603071
GENERAL INFORMATION:
APPLICANT: KILDS.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAN NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Elloppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/NS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CLASSIFICATION: 435
ATORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
TELECOMMUNICATION INPORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: 23: SEQUENCE CHARACTERISTICS:
```

ö

Gaps

ö

Length 28; Indels

g

ô

```
NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEBATITIS C VIRUS
NUT THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0%; Score 151; DB 2;
100.0%; Pred. No. 2.6e-14;
ilve 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLECTIDE AND DEDUCTION OF INVENTION: CORE GENES OF ISOLATION: CORE GENES OF ISOLATION: CORE GENES OF ISOLATION: SEQUENCES IN DIAGNOS INUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS: 263

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                      CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-4UG-1994
CLASSIFICATION: 435
ATTORNEY AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: US/08/290,665A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 190, Application US/08290665A Patent No. 5882852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: US'
ZIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
                      NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEX: 421792
INFORMATION FOR SEO ID NO: 188:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S!
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OPERATING SYSTEM:
                                                                                                                                                  CITY: NEW YORK STATE: NEW YORK
                                                                                                                                                                                                                   USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-290-665A-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-290-665A-190
                                                                                                                                                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                          GENERAL INFORMATION:

APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: BUKELL, R.H.

APPLICANT: BUKELL, R.H.

TITLE OF INVENTION: NUCLECTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: AND THE USE OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ouery Match 100.0%; Score 151; DB 2; Best Local Similarity 100.0%; Pred. No. 2.6e-14; Matches 28; Conservative 0; Mismatches 0;
                      1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPORKTKRNTIRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: FLOPPY DISK
COMPUTER: IMP PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MILLER, R.H. AND
                                                                                                                                                                                                   Sequence 187, Application US/08290665A
Patent No. 5882852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 188, Application US/08290665A
Patent No. 5882852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION: 435
ATORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-6849
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: homosapiens
INDIVIDUAL ISOLATE: HK10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         : 345 PARK AVENUE
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 191 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
APPLICANT: BUKH, J.,
APPLICANT: PURCELL, RITLE OF INVENTION: N
TITLE OF INVENTION: A
TITLE OF INVENTION: C
TITLE OF INVENTION: C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGINAL SOURCE:
                                                                                                                                                                     US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 3
US-08-290-665A-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
```

ద

ô

ö

Gaps

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: CORE GENES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF THE PATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 191;
                                                                                                                                                                                                                          Ouery Match
100.0%; Score 151; DB 5;
Best Local Similarity 100.0%; Pred. No. 2.6e-14;
Matches 28; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 151;
                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRICE APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRICR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 188, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0%;
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLLOGY: unknown
ORIGINAL SOURCE:
                                                                                                                                     ORGANISM: homosapiens : INDIVIDUAL ISOLATE: HK10 PCT-US95-10398-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INDIVIDUAL ISOLATE: S52
PCT-US95-10398-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: unknown ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: amino acid
STRANDEDNESS: unl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ZIP: 10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                       PCT-US95-10398-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE OF INVENTION: UNCLEORIDE AND DEDUCED
TITLE OF INVENTION: ORE GENES OF ISOLATES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEATITIES C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
TOWNERS OF SEQUENCES:
ADDRESSE: MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 100.0%; Pred. No. 2.6e-14;
Matches 28; Conservative 0; Mismatches 0;
                                     ATTORNEY/AGENT INPORMATION:
NAME: RICHARD W. BORK
RECISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEX: 421792
INFORMATION FOR SEQ ID NO: 190:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 anino acids
TYPE: amino acid
STRANBEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUPUTER TOTAL
COUPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION DATA:
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 187, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO: 187:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (212) 758-4800
(212) 751-6849
    FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: homosapiens
INDIVIDUAL ISOLATE: DK12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S 345 PARK AVENUE
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                  -08-290-665A-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PCT-US95-10398-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
```

ò

```
LIEVEN
NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 230, Application US/08635886C
Sequence 230, Application US/08635886C
Patent No. 6555114
GENERAL INFORMATION:
APPLICANT: LERCOUX-NOELS, Geert
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: URMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VIRUS
FILE REFERENCE: 2752-18
CURRENT APPLICATION NUMBER: US/08/635,886C
CURRENT FILING DATE: 1994-10-28
PRIOR APPLICATION NUMBER: PCT/EP94/03555
PRIOR FILING DATE: 1994-10-28
PRIOR FILING DATE: 1993-11-04
NUMBER: OF SEQ. DID NOS! 286
NUMBER: OF SEQ. DID NOS! 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                   STREET: P.O. BOX 4433
CITY: BOX 4020N
STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIDLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MACOSOFT WORD 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,075A
FILING DATE: 21 Apr 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/04155
FILING DATE: 23 Oct 1995
FILING DATE: 21 Oct 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 23 Out 1995
FILING DATE: 22 Jun 1995
ATPORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
RESTERRENCE/DOCKET NUMBER: 29,775
REFERENCE/DOCKET NUMBER: 29,775
REFERENCE/DOCKET NUMBER: 29,775
REFERENCE/DOCKET NUMBER: 29,775
REFERENCE/DOCKET NUMBER: 100: 44:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 151;
Pred. No. 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPORKTKRNTIRRPODVKFPGGGGIVG 32
                                                                                                                                                                                                              ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433 CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0%;
ilarity 100.0%;
Conservative 0;
                                                                                                                                                  AGENTS
207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS:
LENGTH: 319 amino acids
TYPE: amino acid
                 GENERAL INFORMATION:
APPLICANT: MAERTENS, GEE;
APPLICANT: MAERTENS, GEE;
TITLE OF INVENTION: NEW
TITLE OF INVENTION: AND;
TITLE OF INVENTION: AND;
TUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches 28; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-635-886C-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-836-075A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                             Gaps
                                                                                                                                                                                                                                                                                                                                                           AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEBATITIE C VIRUS MAID THE USE OF REAGENYE DERIVED FROM THESE SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100.0%; Score 151; DB 5; Length 191; 100.0%; Pred. No. 2.6e-14; Live 0; Mismatches 0; Indels
                           Indels
100.0%; Pred. No. 2.6e-14;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                     NUCLEOTIDE AND DEDUCED
                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                              APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RIGHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 30,459
REFERENCE/DOCKET NUMBER: 30,459
REFERENCE/DOCKET NUMBER: 2026-4116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                 APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                     RESULT 7
PCT-US95-10398-190
; Sequence 190, Application PC/TUS9510398
; Septence 190, Application PC/TUS9510398
; APPLICANT: BUKH, J., MILLER, R.H. AN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 8
US-08-836-075A-44
; Sequence 44, Application US/08836075A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0%; F1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEPHONE: (212) 758-4800
TELEX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 190: SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: anino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM: homosapiens
INDIVIDUAL ISOLATE: DK12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : 345 PARK AVENUE
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 100.0
Best Local Similarity 100.0
Matches 28; Conservative
Best Local Similarity 100.0
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                TITLE OF INVENTION: MUCL
TITLE OF INVENTION: AAIN
TITLE OF INVENTION: AND
TITLE OF INVENTION: AND
TITLE OF INVENTION: SEQUENCES: 263
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS: 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM: MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PCT-US95-10398-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                  g
```

Gaps

PatentIn version 3.1

ò a

```
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: MUCLECTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 99.3%; Score 150; DB 2; Best Local Similarity 96.4%; Pred. No. 3.6e-14; Matches 27; Conservative 1; Mismatches 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPORKTKRNTIRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                            ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: ALL. PC-DOS/MS COPERATING SYSTEM: PC-DOS/MS COFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
APPLICATION NUMBER: PCT/US95/10398
                                                                                                               SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 189, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION UNMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                            PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO: 189: SEQUENCE CHARACTERISTICS: LENGTH: 191 amino acids TYPE: amino acid STRANDEDNESS: unknown TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3: MORGAN & FINNEGAN
345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDIUM TYPE: FLOPPY DISK COMPUTER: IBM PC COMPATIBLE
                                                                    COMPUTER: IBM PC COMPATIBLE
                                             FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: homosaplens; INDIVIDUAL ISOLATE: S2 US-290-665A-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN &
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                       OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NEW YORK
: USA
                                                                                                                                                                                      FILING DATE: 15
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PCT-US95-10398-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AND THE USE OF REAGENTS DERIVED FROM THESE SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: BURCELL, R.H.
APPLICANT: BURCELL, R.H.
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: MUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AND THE USE OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100.0%; Score 151; DB 4; 100.0%; Pred. No. 4.5e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME/KEY: MISC_FEATURE
LOCATION: (156)..(157)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME/KEY: MISC_FEATURE
LOCATION: (174)..(174)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEY: MISC_FEATURE
LOCATION: (177)..(177)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEY: MISC_FEATURE
COCATION: (233)
COTHER INFORMATION: Xaa is any amino acid
US-08-635-886C-230
                                                                                                                                                                                                                                                                                                                                                                                                COCATION: (161)..(161)

THER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCATION: (167)..(167).
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCATION: (171)..(172)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LOCATION: (232)..(232)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                         LOCATION: (144)..(144)
JTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                              LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 189, Application US/08290665A Patent No. 5882852 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE CITY: NEW YORK
                       IYPE: PRT
ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURE:
NAME/KEY: MISC_FEATURE
                                                                                               NAME/KEY: MISC_FEATURE
                                                                                                                                                                                         NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 10
US-08-290-665A-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY:
                                                                                                                      COCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURE:
                                                                                                                                                                 FEATURE:
```

g

```
GENERAL INFORMATION:
APPLICANT: LEROUX-ROELS, Geert
APPLICANT: LEROUX-ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: URMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VIRUS
TITLE OF INVENTION: VIRUS
FILE REFERENCE: 2752-18
CURRENT FILING DATE: 1996-04-25
PRIOR PELLOATION NUMBER: PCT/EP94/03555
PRIOR PELLOATION NUMBER: PP 93402718.6
PRIOR FILING DATE: 1993-11-04
SOFTWARE: PATENT DATE: 1993-11-04
SOFTWARE: PATENT NUMBER: PT 9360 ID NOS: 286
SOFTWARE: PATENT NUMBER: PT 9400718.6
SOFTWARE: PATENT NUMBER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
APPLICANT: LEROIX-ROELS, Geert
APPLICANT: LEROIX-ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VIRUS
FILE REFERENCE: 2752-18
CURRENT APPLICATION NUMBER: US/08/635,886C
CURRENT FILING DATE: 1996-04-25
PRIOR FILING DATE: 1994-10-28
PRIOR PLICATION NUMBER: PP 93402718.6
PRIOR PLICATION NUMBER: EP 93402718.6
PRIOR PLING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: Patentin Version 3.1
SEQ ID NO 192
LENGTH: 450
                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 450;
                                                 Score 146; DB 3; Length 28;
Pred. No. 1.7e-14;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 146; DB 4; Length 45
Pred. No. 3.4e-13;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 96.7%; Score 146; DB 4; Best Local Similarity 96.4%; Pred. No. 3.4e-13; Matches 27; Conservative 0; Mismatches 1
                                                                                                                                                                    1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                   1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                        Sequence 191, Application US/08635886C Patent No. 6555114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 192, Application US/08635886C
Patent No. 6555114
                                              Ouery Match
Best Local Similarity 96.4%;
Matches 27; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  96.7%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; ORGANISM: hepatitis C virus US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 96.7
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                               JS-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: PRT
                                                                                                                                                                       ò
                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 25, Application US/08921887
Patent No. 603071
GENERAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 150; DB 5; Length 191;
Pred. No. 3.6e-14;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: USA

ZIP: 30303-1769

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E: JONES & ASKEW, LLP
191 Peachtree Street, N.W., 37th Floor
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/POCKET NUMBER: 2026-4116
FELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELERAX: (212) 751-6849
FELERAX: 421792
INFORMATION FOR SEQ ID NO: 189:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     36,714
FR: 03063-0380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE/DOCKET NUMBER: 03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SS: not relevant
not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 25:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99.3%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; ORGANISM: Hepatitis virus US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,7
                                                                                                                                                                                                                                                                     LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 96.4*
Local 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: homosapiens INDIVIDUAL ISOLATE: S2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: unknown
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
STRANDEDNESS: no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ANTI-SENSE: NO
FRAGMENT TYPE: 1
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT-US95-10398-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ξ
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                  APPLICANT: BURCELL, R.H.
APPLICANT: BURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match

96.0%; Score 145; DB 2; Length 191;
Best Local Similarity 96.4%; Pred. No. 1.9e-13;
Matches 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IN PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-ANG-1994
CLASSIFICATION NUMBER: 2026/116
TLING DATE: 15-ANG-1994
CLASSIFICATION NUMBER: 2026-4116
FILING DATE: 15-ANG-1994
CLASSIFICATION NUMBER: 2026-4116
FILING DATE: 15-ANG-1994
CLASSIFICATION NUMBER: 2026-4116
FILEPHONE: (212) 751-6849
FILEPHONE: (212) 751-6849
FILEFEX: 421792
INFORMATION FOR SEQ ID NO: 175:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDRESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                  Sequence 175, Application US/08290665A; Patent No. 5882852; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: homosapiens INDIVIDUAL ISOLATE: P8
                                                                                                               RESULT 15
US-08-290-665A-175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-290-665A-175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
      ô
                                           q
```

Search completed: August 7, 2003, 11:23:50 Job time : 11.6364 secs

a

us-09-491-146a-23.rapb

```
August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Cgn2_6/ptodata///pubpaa/US07_PUBCOMB.ppp:*

(cgn2_6/ptodata///pubpaa/US07_NEW_PUB.ppp:*

(cgn2_6/ptodata///pubpaa/US06_NEW_PUB.ppp:*

(cgn2_6/ptodata///pubpaa/US06_PUBCOMB.ppp:*

(cgn2_6/ptodata///pubpaa/US07_NEW_PUB.ppp:*

(cgn2_6/ptodata///pubpaa/US08_NEW_PUB.ppp:*

(cgn2_6/ptodata///pubpaa/US08_NEW_PUB.ppp:*

(cgn2_6/ptodata///pubpaa/US08_PUBCOMB.ppp:*

(cgn2_6/ptodata///pubpaa/US08_PUBCOMB.ppp:*

(cgn2_6/ptodata///pubpaa/US08_PUBCOMB.ppp:*

(cgn2_6/ptodata///pubpaa/US09_PUBCOMB.ppp:*

(cgn2_6/ptodata///pubpaa/US09_PUBCOMB.ppp:*

(cgn2_6/ptodata///pubpaa/US09_PUBCOMB.ppp:*

(cgn2_6/ptodata///pubpaa/US09_PUBCOMB.ppp:*

(cgn2_6/ptodata///pubpaa/US10A_PUBCOMB.ppp:*

(cgn2_6/ptodata///pubpaa/US10A_PUBCOMB.ppp:*

(cgn2_6/ptodata///pubpaa/US10A_PUBCOMB.ppp:*

(cgn2_6/ptodata///pubpaa/US10_NEW_PUB.ppp:*

(cgn2_6/ptodata///p
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           451899
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-491-146A-23
151
1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          451899 seqs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                           OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Searched:
                                                                                                                                                                                                                                                                                                                                 Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

		Description	Seguence 44. Appl	Sequence 1, Appli	Sequence 10, Appl	Sequence 1, Appli		Sequence 77, Appl	4	Sequence 78, Appl	ø	Sequence 60, Appl	15,	15,	42, A	Sequence 44, Appl	Sequence 42, Appl	
		ID	US-09-851-138-44	US-10-367-677-1	US-09-851-138-10	US-09-758-308-1	US-09-756-875-8	US-09-921-397-77	US-09-851-138-14	US-09-921-397-78	US-09-851-138-46	US-09-851-138-60	US-09-899-046-152	US-09-878-281-152	US-09-899-046-42	US-09-899-046-44	US-09-878-281-42	
		EQ.	10	15	10	6	10	10	10	10	10	10	11	11	11	11	11	
	Query	Length	319	44	74	91	97	103	108	113	137	138	166	166	169	169	169	
æ	Query	Match	100.0	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	
		Score	151	144	144	144	144	144	144	144	144	144	144	144	144	144	144	
-	Result	No.		10	m	4	5	9	7	80	6	10	11	12	13	14	15	

Sequence 44, Appl Sequence 2, Appl Sequence 3, Appl Sequence 76, Appl Sequence 12, Appl Sequence 18, Appl Sequence 50, Appl Sequence 51, Appl Sequence 50, Appl Sequence 51, Appl Sequence 51, Appl Sequence 51, Appl Sequence 23, Appl Sequence 23, Appl Sequence 24, Appl Sequence 20, A	rt output
11 US-09-878-281-44 12 10 US-09-929-955-2 13 11 US-09-929-955-2 14 US-09-194-949-3 15 10 US-09-194-949-3 16 10 US-09-851-138-15 19 10 US-09-851-138-15 19 11 US-09-899-046-52 19 11 US-09-899-046-52 19 11 US-09-899-046-54 19 11 US-09-899-046-50 10 US-09-978-281-50 10 US-09-978-281-50 10 US-09-978-281-50 10 US-09-978-281-50 11 US-09-995-275-29 11 US-09-995-275-29 11 US-09-995-275-29 11 US-09-995-891-894-3 11 US-09-91-81-84-3 11	TEM: PC-OMPALIDIDE  TOSOÍT WOLD 6.0 / ASCII text NN DATA: UNMBER: US/09/851,138 09-MAY-2001 DATA: UNMBER: 08/836,075 CUN KNOOWN- UNMBER: EP 94870166.9 21 Oct 1994 UNMBER: EP 95870076.7 28 Jun 1995 ORMATION: R, PATRICIA A. RUMBER: 29,775 KET NUMBER: 1NNS:004
144 95.4 169 11 144 95.4 182 10 144 95.4 182 10 144 95.4 191 11 144 95.4 319 10 144 95.4 319 11 144 95.4 310 11 144 95.4 3011 10 144 95.4 3	TEK: 15M PC TING SYSTEM: ARE: MICTOSO PLICATION DAP TATION NUMBE: S DATE: 09-M TATION NUMBE: ATION NUMBE: ANGE/DOCKET SEQ ID NO:
16 144 995 19 144 995 19 144 995 20 144 995 22 144 995 23 144 995 24 144 995 25 144 995 31 144 995 33 144 995 33 144 995 34 144 995 34 144 995 35 144 995 36 144 995 37 144 995 38 144 995 38 144 995 39 144 995 30 144 995 31 144 995 32 144 995 34 144 995 35 144 995 36 144 995 37 144 995 38 144 995 38 144 995 39 144 995 39 144 995 30 144 995 31 144 995 31 144 995 32 144 995 33 144 995 34 144 995 35 144 995 36 144 995 37 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995 38 144 995	COMPOTER: OPERATING SOFTWARE: CURRENT APPLICATIC FILING DAT PRIOR APPLICATIC FILING DAT APPLICATIC RELISTRATI
	DANI

```
Sequence 1, Application US/09758308
Patent No. US200200906071
GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C \( \) TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 14114.034902
CURRENT APPLICATION NUMBER: US/09/758,308
PRIOR FILING DATE: 2001-01-10
PRIOR FILING DATE: 60/092,339
PRIOR FILING DATE: 1999-07-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin version 3.0
ENGTHARE: Patentin version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                        CURRENTING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 6.0 / ASCII text output
CURRENT APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 OCt 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 21 OCt 1994
APPLICATION NUMBER: EP 95870076.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          95.4%; Score 144; DB 10; 96.4%; Pred. No. 2.2e-13; vative 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 144; DB 9;
Pred. No. 2.8e-13;
                                                                WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: KAMMERER, PATRICIA A.
RESISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                  ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
0
                                                                                         STREET: P.O. BOX 4433
CITY: HOUSTON
   NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        95.4%;
ilarity 96.4%;
Conservative
                                                                ADDRESSEE: ARNOLD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: amino acid
                                                                                                                                                   STATE: TEXAS COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 27; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAL INFORMATION:
APPLICANT: MERTENS, GEBRT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: LACOUX, AVAIER
TILLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILE REFERENCE: 103959
CURRENT APPLICATION NUMBER: US/10/367,677
CURRENT APPLICATION NUMBER: US/209/389,756
PRIOR APPLICATION NUMBER: US/209/389,756
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-09-07.
RIOR ELING DATE: EARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SEQ ID NO: 2.1
FROM: PLANCE PARENTING AND USES IN THE PRIOR APPLICATION OF USES IN THE PRIOR APPLICATION OF USES IN THE PARENTIAL OF USES IN T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                      Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 44;
                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                      DB 10;
                                                                                                                                                                                                                             Ouery Match
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 28; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 144; DB 15;
Pred. No. 1.3e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                 ; TOPOLOGY: linear; MOLECULE TYPE: peptide; SEQUENCE DESCRIPTION: SEQ ID NO: 44: US-09-851-138-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PKPORKTKRNTIRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 10, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
                           LENGTH: 319 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 1, Application US/10367677 ublication No. US20030118604A1 ENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: JOLIVET, MICHEL
APPLICANT: PENIN, FRANCOIS
APPLICANT: DALBON, PASCAL
APPLICANT: LADAVIERE, LAURENT
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95.4%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: Hepatitis C virus PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 95.4%
Best Local Similarity 96.4%
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DATE: 1991
; RELEVANT RESIDUES: 2 TO 45
US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE: of Hepatitis Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PAGES: 3392-3396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 2
US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
Sequence 14, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
STUTYPER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CORRESPONDENCE ADDRESS: 20,
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER: READALE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRICATION NUMBER: 09/836,075
FILING DATE: CURROWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
RILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 108;
                                                                    Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 144; DB 10;
Pred. No. 3.3e-13;
                                                               Score 144; DB 10;
Pred. No. 3.2e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                        18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              . Sequence 78, Application US/09921397
; Patent No. US20020151484A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 95.4%;
Best Local Similarity 96.4%;
Matches 27; Conservative (
                                                                  Query Match 95.4%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS
; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-851-138-14
                                                                                                                                                                                                                                                                                                         US-09-851-138-14
                                                                                                                                                                                                                                                                              RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 77, Application US/09921397

Sequence 77, Application US/09921397

Patent No. US20020151484A1

GENERAL INFORMATION:
APPLICANT: HYBRIGENICS

TITLE OF INVEWTION: Dathogenic strain of the hepatitis C virus and
TITLE OF INVEWTION: applications thereof
FILE REFRENCE: B409A - JAZ

CURRENT APPLICATION NUMBER: US/09/921,397

CURRENT FILING DATE: 2001-08-02

PRIOR FILING DATE: 2000-08-03

NUMBER OF SEQ ID NOS: 156

SOFWARE: PATENTIN Ver. 2.1

SEQ ID NO 77

TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 97;
                                                                         us-uy-/Jo-u/Jo-u

sequence 8, Application US/09756875;

Patent No. US20020150990A1

GENERAL INFORMATION:

APPLICANT: PIKE, IAN

TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES

NUMBER OF SEQUENCES: 29

CORRESPONDENCE S. 29

CORRESPONDENCE S. 29

CORRESPONDENCE S. 29

CORRESPONDENCE S. 29

COUNTRY STREET: 555 13th Street, N. W. W. STATE:

STATE: D. C. C.

COUNTRY: U. S. C.

COUNTRY: U. S. C.

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC COMPATIBLE

COMPUTER: IBM PC COMPATIBLE

OPPRATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PATENTIN PC-DOS/MS-DOS

SOFTWARE: PATENTIN DATA:

APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 95.4%; Score 144; DB 10
Best Local Similarity 96.4%; Pred. No. 3e-13;
Matches 27; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-AUG-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G.
REFIRENCE/DOCKET NUMBER: 1808-157A
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: peptide US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-921-397-77
                                                                    JS-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

us-09-491-146a-23.rapb

ö

```
APPLICANT: MAERTENS, GEERT
APPLICANT: MAERTENS, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICLOSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                          7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 144; DB 10;
Pred. No. 4.4e-13;
); Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
FILING DATE: CURNOWD
FILING DATE: CURNOWD
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                 WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                       5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 152, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                             Sequence 60, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 138 amino acids
                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                       NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match
Best Local Similarity 96.4
Matches 27; Conservative
                                       27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
US-09-899-046-152
                                                                                                                                                                                                           US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-851-138-60
                                       Matches
                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STUTVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                     polypeptides selected from a
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                            Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                        Score 144; DB 10;
Pred. No. 3.5e-13;
); Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 144; DB 10;
Pred. No. 4.3e-13;
; TITLE OF INVENTION: SID nucleic acids and pol
; TITLE OF INVENTION: pathogenic strain of the
; TITLE OF INVENTION: applications thereof
; CURRENT APPLICATION NUMBER: US/09/921,397
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: EP 00402225.7
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: Patentin Ver. 2.1
; EXO ID NO 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
PAPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE/DOCKET NUMBER: INNS:004 INFORMATION FOR SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPORKTKRNTIRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 46, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH: 137 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD,
                                                                                                                                                                                                                                                                                    TYPE: PRT; ORGANISM: Hepatitis C virus
US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 9
US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3
```

us-09-491-146a-23.rapb

```
Floppy disk
                                                                                                                                                                                                                                                                            linear
       MEDIUM TYPE:
                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                        TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New sequences of hepatitis C virus
genotypes for diagnosis, prophylaxis and therapy.
270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/NS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylax NUMBER OF SEQUENCES: 270
COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 144; DB 11;
Pred. No. 5.3e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                DB 11;
                                                                                                                                                                                                                                                                                                                              Query Match 95.4%; Score 144; DB 11; Best Local Similarity 96.4%; Pred. No. 5.3e-13; Matches 27; Conservative 0; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                            PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 152, Application US/09878281
Publication No. US20030032005A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : Sequence 42, Application US/09899046; Publication No. US20030008274al; GENERAL INFORMATION:
                                                                                                               PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                 FILING DATE:
INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILLING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT:
TITLE OF INVENTION: New
TITLE OF INVENTION: gen
NUMBER OF SEQUENCES: 27C
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                              TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               , TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 27; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270\,
                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0
                                                                                                                                                                                                                                                                                                                                                          Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                   Indels
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 144; DB 11;
Pred. No. 5.4e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                    Score 144; DB 11;
Pred. No. 5.4e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                         PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 44, Application US/09899046; Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 42, Application US/09878281; Publication No. US20030032005A1 GENERAL INFORMATION:
APPLICANT: TITLE OF INVENTION: New sequence: TITLE OF INVENTION: qenotypes for NUMBER OF SEQUENCES: 270
                                                                                                                                                                                   INFORMATION FOR SEQ ID NO: 42: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INFORMATION FOR SEQ ID NO: 44: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                        95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          95.4%;
                                                                                                                                                                                                                            LENGTH: 169 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                : 169 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 96.4'
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 95.48
Best Local Similarity 96.48
Matches 27; Conservative
                                                                                                                                                                                                                                                                                 MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-899-046-44
```

ö

```
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

CONTENTENT ISM PC comparatible

CORRENT APPLICATION DATA:

CURRENT APPLICATION DATA:

PRICATION DATA:

PRILING DATE:

INFORMATION FOR SEQ ID NO: 42:

INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:

LENGTH: 169 amino acids

TYPE: amino acid

TOPOLOGY: 1 inear

MOLECULE TYPE: protein

US-09-878-281-42

OUGLY MAtch

OUGLY MAtch

ON: 5.46-13;

atches 27; CONSETVATION PREGGGOING 28

| INFORMATION PROPER CHARACTERISTICS:

| APPLICATION NO: 5.46-13;

| APPLICATION OF AUGUST PROBUME PROGGGOING 28

| APPLICATION OF AUGUST PROBUME PROGGOING 28

| APPLICATION OF AUGUST PROBUME PROGGOING 28

| INFORMATION PROBUME PROGGGOING 28

| INFORMATION PROBUME PROGGGOING 28

| INFORMATION PROBUME PROGGGOING 32
```

Search completed: August 7, 2003, 12:01:12 Job time : 15.3636 secs

us-09-491-146a-24.rapb

```
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 108;
                                                       Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IEM PC compatible
COMPUTER: IEM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 149; DB 10;
Pred. No. 1.9e-13;
1; Mismatches 0;
                                                     Score 149; DB 10;
Pred. No. 1.8e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
NAME: KAMBERE, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                            1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                            WHITE & DURKEE
                                                                                                                                                   18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
                                                                                                                               1 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STREET: P.O. BOX 4433
CITY: HOUSTON
STRIE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                               Sequence 14, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 78, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                        APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                 AGENTS
                                                       97.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 97.4%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                    CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD,
; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                 Query Match
Best Local Similarity 96.4<sup>1</sup>
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 8
US-09-921-397-78
                                                                                                                                                                                                                                             JS-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-851-138-14
                                                                                                                                 ò
                                                                                                                                                                   윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polypeptides selected from a
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 97;
                                                                                                                                                                                                                                                                                                                                                                                         SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 149; DB 10;
Pred. No. 1.7e-13;
1; Mismatches 0;
                                                                                                    GENERAL INFORMATION:
APPLICANT: PIKE, IAN
TILLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sulte 701-E Columbia Square
STREET: 555 13th Street, N. W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: HYBRIGENICS
TITLE OF INVENTION: 510 nucleic acids and printe OF INVENTION: 510 nucleic acids and printe OF INVENTION: pathogenic strain of TITLE OF INVENTION: applications thereof FILE REFERENCE: B48094 - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILLING DATE: 2001-08-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPOROTKRNTNRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION NUMBER: 2001-08-0225.7
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 77
LENTH: 103
                                                                                                                                                                                                                                                                                           ZIP: 2004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
CMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 77, Application US/09921397 Patent No. US20020151484A1
                                                                Sequence 8, Application US/09756875
Patent No. US20020150990A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-756-875-8
                                                                                                                                                                                                                                                               D. C.
                                                                                                                                                                                                                                                         STATE: D
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
```

```
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPEXLACTIC, THERAPEUTIC AND DIAGNOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: TEAAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER: USA
COMPUTER: FORM:
MEDIUM TYPE: Floppy disk
COMFUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFWARE: Microsoft Word 6.0 / ASCII text output
APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: CUNKnown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 149; DB 10;
Pred. No. 2.5e-13;
                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ANNOLD, WHITE & DURKEE
STREET: P.O. BOX 4443
            Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: KAMBERE, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mismatches
                                                   1 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                          1 PKPOROTKRNTNRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 152, Application US/09899046 Publication No. US20030008274A1
                                                                                                                                                                                            Sequence 60, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 138 amino acids
            ä
                                                                                                                                                                                                                                                          APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                            AGENTS
                                                                                                                                                                                                                                                                                STUYVER, LIEVEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INFORMATION FOR SEQ ID NO: 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 96.49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: amino acid
          27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                              CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                             STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
US-09-899-046-152
                                                                                                                                                     RESULT 10
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-851-138-60
            Matches
                                                   ò
                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: MAERTENS, GEERT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                Gaps
        polypeptides selected from
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER: FLOPPY disk
COMPUTER: FLOPPY disk
COMPUTER: FLOPPY disk
COMPUTER: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION DATA:
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                       Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 137;
                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                     DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 149; DB 10;
Pred. No. 2.5e-13;
                                                                                                                                                                                                                                                                                                                                   Score 149; DB 10
Pred. No. 2e-13;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
TITLE OF INVENTION: SID nucleic acids and portion of INVENTION: pathogenic strain of the TITLE OF INVENTION: applications thereof FILE REFERENCE: B4809A - JAZ CURRENT APPLICATION NUMBER: US/09/921,397 CURRENT FILING DATE: 2001-08-02 PRIOR APPLICATION NUMBER: EP 00402225.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: INNS:004 INFORMATION FOR SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear;
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-851-138-46; Sequence 46, Application US/09851138; Publication No. US20020183508A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 137 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97.48;
                                                                                                                                                                                                                                                                                                                                   ch 97.4%;
1 Similarity 96.4%;
27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                   ; ORGANISM: Hepatitis C virus US-09-921-397-78
                                                                                                                                                                  NUMBER OF SEQ ID NOS: 156
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
                                                                                                                                                                                                                                   LENGTH:
```

```
TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 169;
                                                                                                                                                                                                                                                                                                                                                          DB 11; Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM FC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; DB 11; 1
3.1e-13;
                                                                                                                                                                                                                                                                                                                                                      Score 149; DB 11;
Pred. No. 3.1e-13;
1; Mismatches C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; Score 149; DB
; Pred. No. 3.1e-
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPORQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 44, Application US/09899046; Publication No. US20030008274A1 GENERAL INFORMATION: APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 42, Application US/09878281; Publication No. US20030032005A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97.48;
                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                       LENGTH: 169 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity 96.49
Watches 27; Conservative
                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                             ; MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; MOLECULE TYPE: protein US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14
US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy. NUMBER OF SEQUENCES: 270 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                   Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 166;
                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
   COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                     DB 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 11;
                                                                                                                                                                                                                                                                                                                                   Score 149; DB 11
Pred. No. 3e-13;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 149; DB
Pred, No. 3e-1
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                PKPOROTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 152, Application US/09878281 Publication No. US20030032005A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 42, Application US/09899046
Publication No. US20030008274A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                    PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                   97.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 97.4%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                    I: 166 amino acids amino acid
                                                                                                                                                          FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acidi
                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                       TOPOLOGY: Ilnear
MOLECULE TYPE: protein
US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                >,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
```

```
us-09-491-146a-24.rapb
```

```
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM Pc compatible
COMPUTER: COMPUTER: Patentin Release #1.0, Version #1.25 (EPO)
CORTWARE APPLICATION DATA:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
FILING DATE:
FILING DATE:
FILING DATE:
FILING DATE:
COMPUTER: CHARACTERISTICS:
CHARACTERIST
```

HCV capsid peptide Non-Y, non-B hepat Hepatitis C virus Hepatitis C virus Human hepatitis C HCV core-envelope HCV core-envelope HCV core-envelope HCV core-envelope HCV core-envelope

AAR30689 AAR54065 AAW37380 AAW66083 AAY26952 AAY94410 AAY94409 Non-A, Non-B Hepat Peptide VIIIE base Anti-HCV antibody Prototype peptide HCV core protein p PT-ANB VIIIE Stru Hepatitis C virus Antigen pHCb101. Antigen pHCb101. Antigen pHCb101. Antigen pHCb101. Antigen pHCb101. Antigen pHCb101.

> AAR69545 AAY15425 AAW01865 AAR12597 AAR96530

Branched peptide Hepatitis C virus Antigenic epitope HCV core protein N Hepatitis C virus Protein encoded by

AAR51389 AAR66619 bait polypepti capsid core pr core-envelope

ALIGNMENTS

AAR29530

Hepatitis C virus HCV bait polypepti Protein encoded by

Hepatitis C virus

```
Search time 38.5455 Seconds (without alignments) 115.301 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                         1. SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2. SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3. SIDSI/gcgdata/geneseq-embl/AA1981.DAT:*
3. SIDSI/gcgdata/geneseq-embl/AA1981.DAT:*
3. SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3. SIDSI/gcgdata/geneseqy-embl/AA1982.DAT:*
3. SIDSI/gcgdata/geneseqy-embl/AA1986.DAT:*
3. SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:*
3. SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:*
3. SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:*
3. SIDSI/gcgdata/geneseqy-embl/AA1990.DAT:*
3. SIDSI/gcgdata/geneseqy-embl/AA1990.DAT:*
4. SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
5. SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
5. SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
6. SIDSI/gcgdata/geneseqy-embl/AA2001.DAT:*
          GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                        tal number of hits satisfying chosen parameters:
                                                                                                                                                                          153
1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                            1107863 seqs, 158726573 residues
                                                                                                 August 7, 2003, 11:05:37;
                                                                                                                                                                                                                                                                                                                                                                   Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                    Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                 A_Geneseq_19Jun03:*
                                                                                                                                                           US-09-491-146A-33
                                                                                                                                                                                                                                                                                                                       Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                    BLOSUM62
                                                                                                                                                                         Perfect score:
Sequence:
                                                                                                                                                                                                                    Scoring table:
                                                                     OM protein
                                                                                                                                                                                                                                                                Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                               Database
                                                                                                   Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		Description	NC mosaic protein	NC mosaic protein	Hepatitis C virus	NC mosaic protein	Blood transmiscibl	HCV-S1 full-length	CN14 fragment of H	HCV capsid peptide	HCV capsid peptide
	;	QI	AAY06683	AAY06678	AAR58593	AAY06675	AAR53417	AAE20477	AAW06487	AAR30687	AAR30688
		e :	20	20	15	50	15	23	16	14	14
	Ouery	Length	28	28	312	28	3010	3010	36	38	38
æ	Ouery	Match	100.0	96.7	92.2	90.3	90.3	90.2	89.5	89.5	89.5
		Score	153	148	141	138	138	138	137	137	137
	Result	. O	-1	7	e	4	2	9	7	<b>8</b> 0	6

RESULT AAY0661	AAYO6683
O X	AAY06683 standard; Protein; 28 AA.
23	AAY06683;
<b>1</b> 5 5 5	17-JUN-1999 (first entry)
S E S	NC mosaic protein amino acid fragment K.
(	Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
SOS	Hepatitis C virus.
X Z	WO9910506-A1.
X E ;	04-MAR-1999.
PF	21-AUG-1998; 98WO-US17385.
. E. S.	25-AUG-1997; 97US-0921887.
4 A 3	(USSH ) US DEPT HEALTH & HUMAN SERVICES.
, L	Fields HA, Khudyakov YE;
X 25 3	WPI; 1999-204671/17.
Y L L	New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for
LΙ	detecting hepatitis infection in an individual

Length 28;

```
artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences AAY06673-683 represent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                     Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                 09-MAY-1995 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 92.9
les 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1994-298800/37.
N-PSDB; AAQ70543.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 312 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                                                                           28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JP06225770-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-JUL-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10-JUL-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16-AUG-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Seguence
                                                                                                                                                                                                                                                                                                                   AAR58593;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                      RESULT 3
 8 \times 8 \times 8 \times 8 \times 8 \times 8 \times 8
                                                                                                                                                                               å
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                  ö
                                       The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The artigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein mosaic gene and protein is also useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The invention relates to a mosaic protein, comprising a plurality of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for
                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                    Length 28;
                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                 Score 153; DB 20;
Pred. No. 2.6e-14;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              peptides from different genotypes of a species detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NC mosaic protein amino acid fragment F.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY06678 standard; Protein; 28 AA.
                                                                                                                                                                                                                                      comprising the NC mosaic protein
                                                                                                                                                                                                                                                                                                 100.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 5; Fig 9; 66pp; English.
            Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                  98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Fields HA, Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1999-204671/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                      28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            17-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             04-MAR-1999
                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY06678;
                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 2
AAY06678
 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAQ70543 is a fragment of hepatits C virus (HCV) or non-A non-B hepatits virus (NANBH) core and envelope one structural regions, encompassing base pairs 349-1285. It codes for AAKS8593 an antigen to a structural region of the HCV virus, which can be used in the diagnosis of NANBH patients and the detection of HCV carriers.
                                                                                                                                                                                                                                                                                                                                                                                                                          Non-A non-B hepatitis virus antigens; NANBH; hepatitis C virus.
                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 312;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A nucleic acid fragment coding Non-A Non-B Hepatitis virus antigens - for diagnosis of NANBH and detection of HCV
                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
Score 148; DB 20;
Pred. No. 1.3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 141; DB 15;
Pred. No. 1.5e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (KOKU-) KOKUSAI SHIYAKU KK.
(SANW) SANWA KAGAKU KENKYUSHO CO.
(SPOEU) JONEN CORP.
(TOKR-) ZH TOKYOTO RINSHO IGAKU SOGO KENKYUSHO.
                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus 349-1285 fragment antigen.
                                          0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1; Mismatches
                                                                                                                           28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2 PKPQRKTKRNANRRPQDVKFPGGGQIVG 29
                                                                                                           1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPORKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                             AAR58593 standard; Protein; 312 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 1; Page 18; 22pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY06675 standard; Protein; 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92JP-0207391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       93JP-0193104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
AAY06675
ID AAY0
```

17-JUN-1999

AAY06675;

WO9910506-A1

04-MAR-1999

21-AUG-1998; 25-AUG-1997;

Claim

```
Leu
                                                                                                                                                                                                         'label- Ser, Arg, Gly
   Location/Qualifiers
                                                                                                                                                                                                                                                                    'label Gln, Pro,
                                                                                                                                                                            /label= Gly, Asp
                             label- His, Arg
                                                                                                                     label- Asp, Asn
                                                                                                                                                                                                                                     label- Thr, Ala
                                                                                                                                                                                                                                                                                                                                                         'label- Asn, Asp
                                                                                                                                                                                                                                                                                                 'label- Lys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             label- Ser, Asn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          abel- Thr, Ala
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         abel- Gly, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       label= Glu, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Tabel- Met, Thr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       label- Asn, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 label- Gly, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             label- Gly, Ile
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /label= Cys, Arg
                                                          'label - Cys,
                                                                                       label- Leu,
                                                                                                                                                  'label- Phe,
                                                                                                                                                                                                                                                                                                                          label Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                               abel- Leu,
                                                                                                                                                                                                                                                                                                                                                                                      label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                   label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             abel= Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  abel- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               label- Val,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     'label= Ala,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    abel= Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 1056
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 1255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 1828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 2143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 2488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 1036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 2144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 2486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 2844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 2862
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 1017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference 1201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 1263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 2968
                                                                                                                                                                                         Misc-difference 398
                                                                                                                                                                                                                                                                                Misc-difference 410
                                                                                                                                                                                                                                                                                                                                        Misc-difference 430
                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 2917
                                                                        Misc-difference 246
                                                                                                    Misc-difference 263
                                                                                                                                 Misc-difference 291
                                                                                                                                                                                                                      Misc-difference 400
                                                                                                                                                                                                                                                                                                            Misc-difference 418
                                                                                                                                                             Misc-difference 31
                                                                                                                                                                                                                                                   Misc-difference 40
                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 47
                                           Misc-difference 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 14
   Key
Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosalc protein is also provided. The method is designated restriction endonuclease assisted rigation (REAL). The mosalc protein and the artificial mosalc protein are useful for detecting a hepatitis infection in an individual. The mosalc gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosalc protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polymerase chain reaction; PCR; amplify; primer; non-A, non-B hepatitis; NANBH; virus; blood transmissible; detection; hepatitis virus; RT-PCR; C100 antibody; HCV RNA; NS5 region.
                                                                              Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     invention relates to a mosaic protein, comprising a plurality of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                              New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           90.2%; Score 138; DB 20; Length 2 89.3%; Pred. No. 3.2e-12; Live 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NC mosaic protein amino acid fragment C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                       (USSH ) US DEPT HEALTH & HUMAN SERVICES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR53417 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Blood transmiscible NANBHV protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                          5; Fig 9; 66pp; English
                                                                                                                                                                                                                 98WO-US17385.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Non-A, non-B hepatitis virus
                                                                                                                                                                                                                                             97US-0921887
                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               17-JAN-1995 (first entry)
                                                                                                                                                                                                                                                                                                      Fields HA, Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Jest Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                  WPI; 1999-204671/17.
                                                                                                                           Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28 AA;
```

Sequence

wery Match

ò ద AAR53417;

RESULT 5

36666666588

g

à

```
The invention relates to nucleic acid construct which comprises an expression cassette including a first polynuclectide region including a 5 non-cooling region (NR) sequence of an RR, virus and at least an N-terminal portion of a coding sequence of RNA virus, a second colynuclectide region including a 3' untranslated region (UTR) sequence of polynuclectide region (UTR) sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus and a third polynuclectide region (UTR) sequence of the virus and a third polynuclectide region conditions; and a promoter sequence being operatively linked to expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette. Nucleic acid construct of the invention is useful for detecting the presence of an RNA virus in a cell. It is also useful for screening anti-viral drugs and determining drug resistance of an RNA virus. The present sequence is Hepatitis C virus (HCV) isolate HCV-SI full-length polyprotein.
 Novel nucleic acid construct useful for detecting the presence of RNA virus, comprises an expression cassette and a promoter operably linked to expression cassette for minus strand RNA transcription of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This sequence is encoded by the oligonucleotide, CN14, and represents the peptide fragment CP14. CP14 is a fragment of the core region of hepatitis C virus (HCV). CP14 may be used in the detection of HCV infection and to raise antibodies against it.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Detection of hepatitis C virus - using oligopeptide fragment of core region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CP14; core region; hepatitis C virus; HCV; detection;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  90.2%; Score 138; DB 23; 89.3%; Pred. No. 4.1e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                            Example 1; Page 70-81; 81pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAW06487 standard; Protein; 36 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CN14 fragment of HCV core region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 1; Page 6; 7pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           93JP-0156026.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              93JP-0156026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity 89.3 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (IMMO ) IMMUNO JAPAN KK.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1995-047903/07
                                                                                                                                                                                                                                                                                                                                                                                                                                 3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  36 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAT45055
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IP06327482-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29-NOV-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                31-JAN-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                 Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAW06487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          antibody.
                                                             cassette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAW06487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                 This sequence is encoded by the genome of a blood transmissible non-A, non-B hepatitis (NANBH) virus. The CDNA sequence was isolated using the primers given in AAQ63500-35. The amplified fragments are used in the detection of hepatitis virus. The target DNA was isolated from serum of chronically infected NANBH patients who were Cl00 antibody-positive and HCV RNA (NS5 region) positive. Reverse transcription-PCR and PCR were performed on CDNA and the total human NANBH DNA was constructed from 23 clones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Nucleic acid construct; expression cassette; non-coding region; NCR; untranslated region; UTR; anti-viral drug; drug resistance; HCV-S1; Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                non-A non-B hepatitis virus DNA - used for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 138; DB 15;
Pred. No. 4.1e-10;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (MOLE-) INST MOLECULAR & CELL BIOLOGY (EHRL/) EHRLICH G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAE20477 standard; Protein; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                    Claim 1; Page 8-20; 22pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3
                   /label- Tyr, Cys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hong
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV-Sl full-length polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          90.2%;
                                                                                                                                                                                                                                                                                             Blood-transmissible non-A nor detection of hepatitis virus
                                                                                                                               92JP-0051885
                                                                                                                                                                   92JP-0051885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24-JUL-2000; 2000US-220248P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 89.33,
-has 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20-JUL-2001; 2001WO-IL00669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-JUL-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Fan YH, Lim SP, Lim
Misc-difference 2990
                                                                                                                                                                                                                                         WPI; 1994-163130/20.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2002-280605/32.
N-PSDB; AAD33038.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3010 AA;
                                                                                                                                                                                                      (KAEN/) KAENNO K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
                                                                                                                                                                                                                                                            N-PSDB; AAQ63499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200208447-A2.
                                                     JP06105690-A
                                                                                                                               10-MAR-1992;
                                                                                                                                                               10-MAR-1992;
                                                                                          19-APR-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31-JAN-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAE20477;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAE20477
ID AAE;
XX
AC AAE;
XX
DT 01-
```

ö

Gaps

Ŋ

Fri Aug

Best Loc Matches

AAR30687 ID AAR: XX AC AAR: XX ZX AAR: DT 25-1 DT 11-1

RESULT 8

a

ð

```
The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 3M and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope artibodies in the diagnosis of non-A, non-B hepatitis (NANDH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                             Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-8 hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       89.5%; Score 137; DB 14; Length 38;
89.3%; Pred. No. 6.1e-12;
ive 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (Updated on 25-MAR-2003 to correct PN fleld.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                       Todd JA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR30689 standard; peptide; 38
                                                                                                                                                                                                                                                                                                                                                                                    (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                 91US-0714471.
91US-0718052.
                                                                                                                                                                                                                                                                                                    92WO-US03635
                                              (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV capsid peptide No. 25.
                                                                                               HCV capsid peptide No. 24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                     Jolley ME, Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1993-018073/02
                                                                                                                                                                                                 Hepatitis C virus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          38 AA;
                                                                                                                                                                                                                                W09222571-A1.
                                                                                                                                                                                                                                                                                                  29-APR-1992;
                                                                                                                                                                                                                                                                                                                                   3-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                   20-JUN-1991;
                                              25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                   23-DEC-1992.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      antibodies.
               AAR30688;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR30689;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR30689
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The sequences given in AAR30665-89 represent fragments of the hepetitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope artibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A,
             Length 36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 38
                                              2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 137; DB 14;
Pred. No. 6.1e-12;
1; Mismatches 2;
            Score 137; DB 16;
Pred. No. 5.8e-12;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Updated on 25-MAR-2003 to correct PN field.)
                                                                                               1 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                           AAR30687 standard; peptide; 38 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR30688 standard; peptide; 38 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (BAXT ) BAXTER DIAGNOSTICS INC.
            89.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    91US-0714471.
91US-0718052.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92WO-US03635
                                                                                                                                                                                                                                                               (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                HCV capsid peptide No. 23
                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1993-018073/02
        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      virus – for immu
non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20-JUN-1991;
                                                                                                                                                                                                                                                              25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                  409222571-A1
                                           25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Jolley ME,
                                                                                                                                                                                                                              AAR30687;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
```

ö

Gaps

ö

AAR30688 ID AAR3

RESULT 9

g

õ

Gaps

```
The present sequence represents a Hepatitis C virus (HCV) protein sequence from the disclosure of the present specification. The present specification describes a chimeric HCV peptide antigen which comprises at least 2 peptide epitope regions from the HCV polypeptide core region, 2 peptide epitope regions from the NS3 region and at least 2 peptide epitope regions from the NS4 region. The antigen binds specifically with an antibody produced by a human infected by HCV. The peptide can detect a wide range of HCV infections with high sensitivity.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New chimeric peptide antigen derived from hepatitis C virus protein - useful for detecting HCV infections
                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus; HCV; chimeric; antigen; detection; core region; epitope; NS3; NS4; infection.
                                                                                                                                                                                                      ö
                                                                                     Hepatitis C virus #4 and #6 genes were isolated (AAQ64068-69). Both genes contain the core, ENV, NS1, NS2 and NS3 regions. A core region fragment is given in AAQ64067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
0
                      non-B hepatitis virus antigen - useful for
                                                                                                                                                                           Length 38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89.5%; Score 137; DB 19; Length 43;
89.3%; Pred. No. 7e-12;
11ve 1; Mismatches 2; Indels
                                                                                                                                                                                                    Indels
                                                                                                                                                                      Score 137; DB 15;
Pred. No. 6.1e-12;
1; Mismatches 2;
                                                                                                                                                                                                                                           1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Disclosure; Page 11; 22pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Page 24; 30pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus C-1 protein 1-43.
                                                                                                                                                                                                                                                                                                                        AAW37380 standard; Protein; 43
                                                                                                                                                                     89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97JP-0027015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    96JP-0024045
                   DNA coding a Non-A, non-B }
detecting HCV within serum
                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                               Ouery Match
Best Local Similarity 89.3°
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 89.3 es 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1998-022248/03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (TOFU ) TONEN CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis c virus
                                                                                                                                           38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                43 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JP09278794-A
                                                                                                                                                                                                                                                                                                                                                                              11-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   39-FEB-1996;
                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                 AAW37380;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loca
Matches
                                                                                                                                                                                                                                                                                                            AAW37380
        à
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                 The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contained in the capsid protein of the virus and themselves contained in arrows. These peptides can be used in immunoassays for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                   Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Non-A, non-B hepatitis virus; NANBHV; hepatitis C virus; HCV; core; ENV; NS1; NS2; NS3; antigen; detection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 137; DB 14; Length 38;
Pred. No. 6.1e-12;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Non-A, non-B hepatitis virus corel region fragment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOKR-) ZH TOKYOTO RINSHO IGAKU SOGO KENKYUSHO
                                                                                                                                                                                                                                                                                                                                                                                                                         (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                         Disclosure, Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SANW ) SANWA KAGAKU KENKYUSHO CO
                                                                                                                                    (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AARS4065 standard; Protein; 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                89.5%;
                                                                   92WO-US03635.
                                                                                           91US-0714471.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92JP-0088140,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                               Leahy DC,
                                                                                                                                                                                          WPI; 1993-018073/02,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1994-205028/25.
N-PSDB; AAQ64067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TONEN CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       38 AA;
                                                                   29-APR-1992;
                                                                                           13-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JP06141870-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12-MAR-1992;
                                       23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       14-FEB-1995
                                                                                                                                                               Jolley ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24-MAY-1994
                                                                                                                                                                                                                                                                                                                                                                                                             antibodies.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR54065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AARS4065
ID AARS
XX
3888888888%8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
```

à

ö

Gaps

```
Monoclonal antibodies useful for detecting and/or quantifying hepatitis \mathsf{C} virus core protein -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polypeptides that bind to anti-hepatitis C virus antibodies, useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                          Peptides AAY26949-Y26955 represent peptide epitopes derived from the N-terminus of the hepatitis C virus core protein. The peptides are used to generate monoclonal antibodies or antibody fragments specific for hepatitis C virus (HCV) core protein. The monoclonal antibodies are used for early diagnosis of HCV infections, especially by two-antibody
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Human; hepatitis C virus; HCV; immunogen; antigen; antibody; virucide; hepatotropic; anti-inflammatory; virus detection; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human hepatitis C virus core protein N-terminus, residues 2-45,
                       Epitope; hepatitis C virus; core protein; moncolonal antibody;
diagnosis; infection; sandwich immunoassay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Jolivet-Reynaud
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 137; DB 20;
Pred. No. 7.1e-12;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                        Paranhos BC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Æ
                                                                                                                                                                                                                                                                                                                                                                                                           Claim 3; Page 10; 19pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY94410 standard; peptide; 44
                                                                                                                                                                                                                                                                                        Yvon S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         89.5%;
                                                                                                                                                                                                                 98FR-0003087.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99WO-IB01933.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98US-0196155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 89.30,
Thes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dalbon
                                                                                                                                                                                                                                                                                      Piga N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sandwich immunoassay
                                                                                                                                                                                                                                                                                                                       WPI; 1999-530397/45
                                                                                                                                                                                                                                                    (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2000-411934/35,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (INMR ) BIO MERIEUX
                                                                          Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        44 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200031130-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            H
                                                                                                           FR2775690-A1
                                                                                                                                                                                                                 09-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11-SEP-2000
                                                                                                                                            10-SEP-1999
                                                                                                                                                                                                                                                                                      Jolivet RC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  02-JUN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dalbon P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY94410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY94410
       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              complex formation). Antibodies are used correspondingly to detect HCV or related antigens. The peptide and antibodies may also be used to treat or prevent HCV infections. The present sequence represents the N-terminus of the core p21 protein of hepatitis C wirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The invention relates to a peptide, which is recognised by antibodies against amino acids 2.45 at the N-terminus of the core (or nucleocapsid) p21 protein of hepatitis c virus (HCV), or its variants. The peptide has a tertiary structure consisting of two alpha-helical fragments, almost perpendicular to each other in space, connected by a junction peptide. Excluded are all proteins and peptides comprising, or consisting of, the N-terminal part of p21 (starting from amino acid 1 or 2). Also new are (1) monoclonal or polyclonal antibodies produced using the peptide as an immunogen and (2) complex consisting of the peptide specifically bound to some other molecule, particularly peptide or nucleotide fragments or functionalised aromatic compounds. The peptide is used (i) quantifying either antibodies against p21 or HCV-deriyed mRNA (by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     nseq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            peptide from the N-terminus of hepatitis C virus p21 protein taining the immunodominant epitope - and related antibodies, udiagnosis, treatment and prevention of hepatitis C infection
                                                                                                                                                                         antigenic; nucleocapsid; p21 protein; alpha-helical; immunogen; antibody; hepatitis C virus; epitope; N-terminal fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     89.5%; Score 137; DB 19; Length 44; 89.3%; Pred. No. 7.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                              Penin F;
                                                                                                                                        Hepatitis C virus p21 protein N-terminal fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus Core protein amino acids 2-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                            Ladaviere L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. 7.16
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKPQRKTKRNTNRRPQDVKFPGGGGIVG 31
                                AAW66083 standard; peptide; 44 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Page 16; 37pp; French
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                            Lacoux X,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY26952 standard; peptide; 44
                                                                                                                                                                                                                                                                                                                                    98WO-FR00442
                                                                                                                                                                                                                                                                                                                                                                     97FR-0002878
                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                          Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                       (INMR ) BIO MERIEUX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1998-495793/42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
tes 25; Conserv
                                                                                                                                                                                                                            hepatitis c virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      44 AA;
                                                                                                    16-NOV-1998
                                                                                                                                                                                                                                                             WO9839360-A1
                                                                                                                                                                                                                                                                                                                                  05-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                     05-MAR-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21-DEC-1999
                                                                                                                                                                                                                                                                                                 11-SEP-1998.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                containing
                                                                                                                                                                                                                                                                                                                                                                                                                                          Dalbon P,
                                                                   AAW66083;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY26952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 14
RESULT 1
AAW66083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY26952

1D AAY2

XX

AC AAY2

XX

DT 21-D

XX

DE Hepa
```

SXSC

ò g

ö

Gaps

ö

ပဲ

Length 44; Indels

Jolivet M;

diagnosing and preventing hepatitis C infections -

Claim 1; Page 42; 50pp; English.

The present peptide, designated 542G, corresponds to residues 2 to 45 of the N-terminus of the human hepatitis C virus (HCV) core protein. It is an immunodominant region containing conformational type epitopes and linear type epitopes. It manifests an immunoreactivity with all the sera of individuals or blood samples infected with HCV and which possess antibodies directed against the core protein. An amino acid may be substituted for homologous amino acids and side chains and peptide bonds may be modified. For example, L-amino acids may be replaced by D-amino acids, amine groups may be acetylated, and so on. The native antigenic sequence and its antigenic derivatives may be used for detection of hepatitis C virus and for raising antibodies against the virus. 

44 AA; Sequence

ô Gaps ; 0 89.5%; Score 137; DB 21; Length 44; 89.3%; Pred. No. 7.1e-12; Live 1; Mismatches 2; Indels Query Match
Best Local Similarity 89.34
Atches 25; Conservative

1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28

ć g

Search completed: August 7, 2003, 11:14:10 Job time : 38.6364 secs

```
GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:
August 7, 2003, 11:05:41; Search time 9.54545 Seconds
(without alignments)
282.095 Million cell updates/sec
Title:
US-09-491-146A-33
Sequence:
1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5
Searched: 283308 seqs, 96168682 residues
tal number of hits satisfying chosen parameters: 283308
```

Maximum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Listing first 45 summaries

Database : PIR\_76:\*

1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	qenome polyprotein																					റ	1	genome polyprotein	genome polyprotein	genome polyprotein	genome polyprotein	hypothetical prote	genome polyprotein
SUMMARIES	ឧ		S41353	S41355	S41357	_	S41371	S41341	S41370	S41369	S41368	S41342	S41344	S41350	S41354	S41345	S41347	S41343	S41346	S21471	S12707	PC1284	JQ1925	JQ1926	JH0711	S19876	\$18031	S18032	PN0677	JQ0883
	Length DB	3010 1			108 2				4	4									118 2											874 2
æ	Query Match I	90.2	89.5	89.5	89.5	89.5	6	6.											. 89.5			•			6	σ.	89.5	6	ó,	-
	Score	138	137	137	137	137	137	137	137	137	137	137	137	137	137	137	137	137	13.7	137	137	137	137	137	137	137	137	137	137	137
	Result No.	-	7	m	4	2	9	7	8	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	53

genome polyprotein polypeptide - hepa genome polyprotein	ALIGNMENTS	(strain JT) protein M; hepacivirin (EC 3.4.21.98) (nons 4b; nonstructural protein NS5 19-May-2000 #text_change 19-Jan-2001	Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hiji	.C virus genome from a single Japanese carr 14; PMID:1318627	; NID:9221612; PIDN:BAA01943.1; PID:922161	5, NG 19; F	Files is systematical envelope process Files predicted cames (1927) 1900-729/Product: nonstructural protein NS1 #status predicted cAS1> Files 1007-106/Product: nonstructural protein NS2 #status predicted cAS2> Files 1007-1615/Product: hepacivirin #status predicted cAS3> Files 1007-1615/Product: nuclectide binding motif A (P-loop) Files 1017/Region: nuclectide binding motif B	in NS4a *status predicted <n4a> in NS4b *status predicted <n4b> in NS5 *status predicted <ns5></ns5></n4b></n4a>	re 138; DB 1; Length 3010; d. No. 2.5e-16; Mismatches 2; Indels 0; Gaps 0;	VG 28         VG 32	(genotype 2, N2) (fragment) 26-Jul-1996 #text_change 17-Nov-2000 wer, J.T.
JQ0881 PC2219 CNWVTC GNWVCT S18030 GNWVCH JC5620 JC5620 JC5620 JC5620 S21336 S21336 S41358 S41358	ALIG	s.C virus (strain envelope protein protein NS4b; non .revision 19-May-:	Σ.	922957	B:D01171;	NCBI bac genome ydrolase C #statu	protein protein protein status Inding m	al prote al prote al prote	Score Pred. 1; Mis	PKPORKTKRKAHRRPODVKFPGGGIVG 	virus .; Brou
000000000000000		s c pro	awa	he	gg :	rus rot	ral ral rin re-b	tur	38 8	0 - 0 - 0	is C e_re. E.M
876 876 876 876 876 876 876 877 877 877	•	epatiti tein C; ctural virus equence	.; Nakagawa,	5573;	GB:D11168;	cted fraction protein protein protein celope protei	or enverstructum structum nepacivi nepacivi nelectid	onstruc ionstruc	90.2% 17 89.3% ervative	(RKAHRRP   :    KRNTYRRP	atit irus 2 uenc r, G
		rotein - hepatit capsid protein C a; nonstructural epatitis C virus ay-2000 #sequenc	(ato, N.	Jy-Jy, lar clc lber: A4 15573 Lminary	COLU CIANY	bepatit	ict: maj lict: nor lict: nor duct: no jion: no jion: no	oduct: r	Similarity 5; Conserva	PKPORKTI           	rot 1s 1s 4 # 8 1
137 137 137 137 137 137 137 133 133 133		pro ca (4a; hep May	a, T.; F	Title: Molecular clo., 71711e: Molecular clo., 7171e: Accession: A45573 (Status: preliminary Molecule type: DNA	ues: I-: -referer imental	sequenc family: rds: ATI /Product	29/Frout 29/Prod 1006/Prod 1615/Pro 1317/Rec	1862/Pro 2013/Pro 3010/Pro	Match Local es 2	ц 53 Н	2 poly ins: es: tty: 19- 19- sion
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~		RESULT 1 A45573 A45573 Genome poly N; Contains: protein NS C; Species: C; Date: 19-	R, Tanak	A; Refer A; Refer A; Acces A; Statu A; Molec	A; Resid A; Cross	A; Note: C; Super C; Keywo F; 2-115	F; 130-7 F; 730-1 F; 1007- F; 1230- F; 1312-	F; 1616- F; 1863- F; 2014-	Query P Best LA Matche	ογ Dp	RESULT 2 S41153 genome poly N;Contains: C;Species: A;Variety: C;Date: 19- C;Accession R;Van Doorn

2

ö

ö

Indels

Pred. No. 1.6e-11;

89.38;

```
Conservative
                                Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Accession: S41348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Accession: S4137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                       депоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           S41341
                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Securine purplements of virus (yenocype 4, nw, (itayment, yenomine) yenome protein inspatitis C virus
A; Variety; genotype 2, N4
C; Species: hepatitis C virus
A; Variety; genotype 2, N4
C; Sacession: S41355
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPa.
A; Reference number: S41341
A; Accession: S41355
A; Molecule type: genomic RNA
A; Residues: 1-108 < VANN
A; Cross-references: EMBL: 229458
A; Experimental source: genotype 2, N4
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
C;Species: hepatitis C virus
A;Variety: genotype 2, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Date: 19-May-1994 #sequence_revision 24-1957
R;Van Doorn, L.J; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Reference number: S41341
A;Accession: S41357
A;Accession
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A.Reference number: $41341
A.Accession: $41343
A.Molecule type: genomic RNA
A.Residues: 1-108 <-WAN
A.Residues: 1-108 <-WAN
A.Residues: 1-108 <-WAN
A.Residues: Coss references: EMBL: 229456
A.Experimental source: genotype 2, N2
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Keywords: capsid protein, core protein, solyprotein
F.1-108/Product: core protein #status predicted <-WAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89.5%; Score 137; DB 2; I
89.3%; Pred. No. 1.6e-11;
tive 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       89.5%; Score 137; DB 2;
89.3%; Pred. No. 1.6e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. 1.6e
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKKKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
Best Local Similarity 89.39
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 89.39
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
```

Length 108;

5 DB

Score 137;

89.58;

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 5, N5
A; Variety: genotype 5, N5
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41371
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
A; Reference to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LipA.
                                                                                                                                                                                                                                                                                                                                             #text_change 17-Nov-2000
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                        C, Accession: 541348
R; van Doorn, L,J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: 541341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                    (genotype 1, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein - hepatitis C virus (genotype 1, N1) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 137; DB 2; Length 112;
Pred. No. 1.6e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A.Experimental source: genotype 1, N6
C.Superfamily: hepatitis C virus genome polyprotein
C.Keywords: capsid protein; core protein; polyprotein
F.1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 137; DB 2; I
Pred. No. 1.6e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A:Experimental source: genotype 5, N5
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; polyprotein
F:1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                  N.Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N6
C;Date: 19-May 1994 #sequence_revision 26-Jul-1996
                                               28
                                                                           1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPORKTKRKAHRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                              polyprotein - hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
A; Cross-references: EMBL:229451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Residues: 1-112 <VAN>
A;Cross-references: EMBL:Z29474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: genomic RNA A; Residues: 1-112 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatit.i
N;Contains: core protein
C;Species: hepatitis C virus
```

```
RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                841368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  541342
                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
A; Variety: genotype 1, N1
C; Date: 19 May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41341
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EmBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41341
A; Rocession: S41341
A; Molecule type: genomic RNA
A; Rolecule type: genomic RNA
A; Residues: 1-112 < VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein - hepatitis C virus (genotype 5, N3) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N3
C;Variety: genotype 5, N3
C;Date: 19-May-1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-2000
C;Accession: S41369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C.Accession: S41370
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Accession: S41341
A; Accession: S41370
A; Residues: 1.114 cVAN>
A; Residues: 1.114 cVAN>
A; Cross_references: EMBL:229473; NID:9443908; PIDN:CAA82611.1; PID:9443909
A; Experimental source: genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: genomic RNA
A; Residues: 1-114 «VAN>
A; Cross-references: EMBL:229472; NID:g443906; PIDN:CAA82610.1; PID:g443907
A; Experimental source: genotype 5, N3
C; Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                 A; Cross-references: EMBL:229444; NID:9443860; PIDN:CAA82582.1; PID:9443851
A; Experimental source: genotypel, NI
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Rivan Doorn, E.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: $41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 137; DB 2; Length 114;
Pred. No. 1.6e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 137; DB 2; Length 112;
Pred. No. 1.6e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Match 89.5%;
Local Similarity 89.3%;
les 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 89.3 Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Accession: S41369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
ö
                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N10
C;Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41342
                                                                                                                                                                                                                                                                                                                                                                   N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 5, N2
C.Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C.Accession: S41368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: genomic RNA
A; Residues: 1-114 <VAN>
A; Cross-references: EMBL: 229471; NID: 9443904; PIDN: CAA82609.1; PID: 9443905
A; Experimental source: genotype 5, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A Wolecule type: genomic RNA
A; Residues: 1-115 < VAN>
A; Residues: 1-115 < VAN>
A; Cross-references: EMBL: Z29445; NID: g443852; PIDN: CAA82583.1; PID: g443853
A; Experimental source: genotype 1, N10
C; Superfamily: hepatitis C virus genome polyprotein
C; Reywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein status predicted < MAT>
                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Lipa
                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (genotype 1, N10) (fragment)
                                                                                                                                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus (genotype 5, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by
A;Reference number: S41341
A;Accession: S41368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        рy
                                                                          Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 115;
                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 137; DB 2; L
Pred. No. 1.7e-11;
1; Mismatches 2;
                                                                     Score 137; DB 2; L
Pred. No. 1.6e-11;
1; Mismatches 2;
C;Keywords: capsid protein; core protein; polyprotein F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 137; DB 2;
Pred. No. 1.6e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. 1.66
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Mismatches
                                                                                                                                                                       28
                                                                                                                                                                                                 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                       1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89.58;
                                                                        89.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                     Query Match 89.5
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Accession: S41342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
```

us-09-491-146a-33.rpr

```
completed: August
ne : 9.54545 secs
A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: S41345
                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          time
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Search o
                                                                                                                                                                                                                                                                                                                                                                                                                                       S41345
                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                     N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N2
C;Date: 19-May 1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N:Contains: core protein
C:Species: hepatitis C virus
A:Variety: genotype 1, N8
C:Date: 19-May 1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus (genotype 2, N3) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N3
C;Date: 19-May_1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Cross-references: EMBL:229453; NID:9443868; PIDN:CAA82591.1; PID:9443869
A; Experimental source: genotype 1, N8
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                             A; Cross-references: EMBL: 229447; NID: 9443856; PIDN: CAA82585.1; PID: 9443857
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                     C, Accession: S4134
C, Accession: S4134
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPa.
A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LiPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
          genome polyprotein - hepatitis C virus (genotype 1, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus (genotype 1, N8) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C; Accession: $41350
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by
A; Reference number: $41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 115,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 137; DB 2; I
Pred. No. 1.7e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             89.5%; Score 137; DB 2; L
llarity 89.3%; Pred. No. 1.7e-11;
Conservative 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                  A: Experimental source: genotype 1, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               uery Match 89.5%;
.est Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                               A; Molecule type: genomic RNA A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: genomic RNA A; Residues: 1-115 < VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: genomic RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Reference number: $41341
A; Accession: $41350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                       A; Accession: S41344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C; Accession: 541354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: S41354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

7, 2003, 11:21:49

```
ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, N3
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Cross-references: EMBL:229448; NID:9443858; PIDN:CAA82586.1; PID:9443859
                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C,Accession: 341345
C,Accession: 341345
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPa.
A;Reference number: 341341
                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö

    hepatitis C virus (genotype 1, N3) (fragment)

                                                                                                                                                               Length 115,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 137; DB 2; Le
Pred. No. 1.7e-11;
"...matches 2;
A)Cross-references: EMBL:229457
A)Experimental source: genotype 2, N3
C)Superfamily: hepatifis C virus genome polyprotein
C)Superfamily: hepatifis C virus genome polyprotein
C)Keywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                             Score 137; DB 2; I
Pred. No. 1.7e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A Experimental source: genotype 1, N3 C. Superfamily: hepatitis C virus genome polyprotein C Keywords: capaid protein; core protein; polyprotein F:1-115/Product: core protein #status predicted KMT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Mismatches
                                                                                                                                                                                                                                                             1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                             1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     89.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: genomic RNA A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
```

7

```
August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                           127863
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                   tal number of hits satisfying chosen parameters:
                                                                                                                                                                                          US-09-491-146A-33
153
1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                              127863 seqs, 47026705 residues
                                                                           OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                      Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                          Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                Searched:
                                                                                                               Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries SwissProt\_41:\* Database :

Maximum DB seq length: 0 Maximum DB seq length: 2000000000

SUMMARIES

	Description	Q00269 h genome po	hepatitis	hepatitis	hepatitis	-	hepatitis	h genome p	P26662 h genome po	h genome	h genome	h genome	h genome	h genome	herpes si	homo sa		_	P03416 murine coro						pongo	bos ta	P40187 saccharomyc	P36417 dictyosteli	Q28193 bos taurus	Q09459 caenorhabdi	s sns	fugn		bos t
•	ΙD		POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLC_HCVJ8	POLG_HCVTW	POLG_HCV1	IE63_HSV2H	FURI_HUMAN	RK4_TOBAC	NCAP_CVM3	NCAP_CVMA5	NCAP_CVMS	RK4_SPIOL	RL29_MYCGA	FURI_MOUSE	ATP8_PONPP	ATP8_PONPY	F262_BOVIN	PIG2_YEAST	GBF_DICDI	FURI_BOVIN		ໝ	s7_	ᆨ	GDF6_BOVIN
	BB	-	-	-	-	٦		7	٦	٦	-	7	-	-	~	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	~	<del></del>	-
	Length	3010	513	520	520	737	737	10	3010	0	03	03	01	0	512	794	282	454	454	454	293	153	793	68	68	230	238	708	79	1531	139	194	194	436
% Query		90.2	9.	6	6	o.	ó,	6	89.5	œ.	ď.	δ.	δ.	4.	4	4.	۳.	e,	m.	۳.	32.7	31.4	ij.	Ö	0	30.7	Ö	ö	ö	ö	ö	o.	30.1	30.1
	Score	138	m	S.	3	m	ß	m	m	3	m	3	m	a	23	53	51	51	51	51	20	48	48	47	47	47	47	47	47	47	46		46	40
Result	No.	7	7	m	4	ഹ	9	7	8	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	53	30	31	32	33

	015530 homo sapien 09z2a0 mus musculu			
APG_BRANA NCAP_CVMJH	PDPK_HUMAN PDPK_MOUSE	KCB1_MOUSE KMLS_BOVIN	PKP4_HUMAN SLAP_BACCI	KALS_HUMAN Y179_HUMAN RS7A_YEAST
20	44.	4 F F F		*06 +44
44	556 559	117	161	181
30.1 30.1	30.1	30.1	30:1	29.7 29.4
4 4 6 6	444	44	444	45.5 45
34 35	36 37	0 6 4 0 0 0	444	44

## ALIGNMENTS

POLG_HCVJT STANDARD;  AC 010-APR-1993 (Rel. 25, Last) DT 01-APR-1993 (Rel. 25, Last) DT 01-APR-1993 (Rel. 25, Last) DT 15-SED-2003 (Rel. 25, Last) DE 01-APR-1993 (Rel. 25, Last) DE 02 01-APR-1993 (Rel. 25, Last) DE 02 02 03 04 02 05 06 06 06 06 06 06 06 06 06 06 06 06 06	HGVUT  POLG_HCVIT  B. (12, Last annotation update)  Genome polyprotein El (GP32) (GP35); Envel  Genome polyprotein El (GP32) (GP35); Envel  GEC 3.4.22); Protease/helicase NS3 (P70) (H  EC 3.4.22); Protease/helicase NS3 (P70) (H  ENDIANCE FROM N.A.  MADLINE-92295714; PubMed=1318627;  HYDROPHORIC, SUGGESTING A POSISIBE MEMBRA  HYDROPHORIC, SUGGESTING A POSISIBE MEMBRA  HYDROPHORIC, SUGGESTING A POSISIBE MEMBRA  NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL  HYDROPHORIC, SUGGESTING A POSISIBE MEMBRA  NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL  HYDROPHORIC SUCCEPROTEIN E. THE NUCLEOC  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOC  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOC  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOC  PROTEIN M AND MNA.	PRT;  1)  1)  1)  1)  1)  1)  1)  1)  1)  1	update update update update update l prote l prote l prote l (P70) l (	PECTION STANDARD; PRT; 3010 AA.  POGGLECUT STANDARD; PRT; 3010 AA.  POGGLECUT STANDARD; Created)  10.1 ARF-1993 (Rel. 25, Created)  10.1 ARF-1993 (Rel. 25, Last sequence update)  10.5 SEP-2003 (Rel. 12, Last annotation update)  6choome Polyprotedia (Contains: Capped protein Center of Core protein NST Potts of Core protein Center of Core protein NST Potts of Center
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                 01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66-OCT-2001 (Rel. 40, Last annotation update)
66-OCT-2001 (Rel. 40, Last annotation update)
67-0CT-2001 (Rel. 40, Last annotation update)
67-0CT-2001 (GP12) (GP12) (GP12);
67-0CT-2001 (GP12) (GP12);
67-0CT-2001 (GP12) (GP12);
67-0CT-2001 (GP12) (GP12);
67-0CT-2001 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genotypes."

Virology 188:331-341(1992).

-!- FUNCTION: THE SAMAL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
-!- FUNCTION: THE SAMAL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
-!- FUNCTION: THE SAMAL PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN MAND GLYCOPROTEIN E. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-92230232; PubMed-1314459; Okada S.I., Yamamoto K., Lizuka H., Okada S.I., Vamamoto K., Lizuka H., Chokanda T., Eukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor monology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R InterPro; IPR00252; HCV_capsid.
R InterPro; IPR00252; HCV_capsid.
R InterPro; IPR00251; HCV_core.
R InterPro; IPR00251; HCV_core.
R InterPro; IPR00251; HCV_core.
R InterPro; IPR00251; HCV_core.
R InterPro; IPR00251; HCV_core; I.
R InterPro; IPR00251; HCV_core; I.
R Pfam; PF01542; HCV_core; I.
R Pfam; PF01560; HCV_NS1; I.
R ProDom; PD186062; HCV_NS1; I.
R POLOPHOTOLEIN; Glycoprotein; Coat protein; Envelope protein; ITRAGSmembrane; Nonstructural protein.
INIT_MET I. I. REMOVED FROM CAPSID PROTEIN C BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC...
                                                                                                                                                                                                                   513 AA
          EMBL; D10074; BAA00968.1; -.
                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                   POLG_HCVJ2
P27959;
                                           ß
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
CHAIN
CHAIN
TRANSMEM
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                              RESULT 2
POLG_HCVJ2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                    CELLULAR ANINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN (POTENTIAL).
NAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ. (POTENTIAL).
PROTEASE/HELICASE NSJ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BY SIMILARITY).
(BY SIMILARITY).
(BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 138; DB 1; Length 3010;
Pred. No. 2e-11;
1; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
CHARGE RELAY SYSTEM (B CHARGE RELA
                                                                                                                                                                                                                                                                                                                  HCV_RdRP.
RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF01542; HCV_COTE; 1. Pfam; PF01539; HCV_COTE; 1. Pfam; PF01560; HCV_NS1; 1. Pfam; PF01560; HCV_NS2; 1. Pfam; PF01006; HCV_NS3; 1. Pfam; PF01006; HCV_NS4; 1. Pfam; PF01001; HCV_NS4; 1. Pfam; PF01506; HCV_NS5s; 1. Pfam; PF00998; Viral_RGRP; 1. Pr0Dom; PD186062; HCV_NS1; 1. Pr0Dom; PD186062; HCV_NS1; 1. Pr0Dom; PD186062; HCV_NS1; 1. SWART; SW00487; DEXDO; 1.
                                                                                                                                                                           HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                        PF01543; HCV_capsid;
PF01542; HCV_core; 1.
PF01539; HCV_env; 1.
Interpro; IPR002522; HG Interpro; IPR002521; HG Interpro; IPR002531; HG Interpro; IPR002531; HG Interpro; IPR002109; HG Interpro; IPR001490; HG Interpro; IPR001490; HG Interpro; IPR002066; HG Interpro; IPR002066; HG Interpro; IPR002066; HG Interpro; IPR007095; Rb Interpro; IPR007094; Rb Interp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        623
645
2041
2077
2240
2529
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
```

ö

Gaps

ö

(POTENTIAL). (POTENTIAL). (POTENTIAL).

(POTENTIAL) (POTENTIAL)

(POTENTIAL). (POTENTIAL).

 $\sim$ 

```
520 5
520 AA;
                                                                                                                                                                                                                                                                              NCBI_TaxID=31644;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Transmembrane;
INIT_MET 1
                                                                                                                                                       POLG_HCVHK
Q01403;
01-JUL-1993 (
01-JUL-1993 (
   NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                     RESULT
   F S
                                                                              ò
                                                                                                  요
                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
BY SIMILARITY.
                                                                                                                                                                                                                                                                                                     Abe K., Inchauspe G., Fujisawa W.;
"Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2729-2739(1992).
I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                             01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last amortation update)
66-OCT-2001 (Rel. 40, Last amortation update)
Envelope 91ycoprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope 91ycoprotein El (GP32) (GP35); Envelope 91ycoprotein E2
(GP68) (GP20) (NS1)] (Fragment).
                                             Gaps
                                                                                                                                                                                                                               Hepatítis C virus (isolate HCV-476) (HCV).
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nonstructural protein.

Nonstructural REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                      Length 513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein; Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
INII_MET 1 1 REMOVED FROM CAPSID PROTEIN C
          Score 137; DB 1; Length or. Pred, No. 4.3e-12;
  943F31E3514CDEF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC.)
                                                                                                                                            520 AA
                                           1; Mismatches
                                                                1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                           EMBL; D10688; BAA01530.1; -.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01559; HCV_env; 1.
Pfam; PF01559; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                  MEDLINE-93019030; PubMed-1383400;
 55704 MW;
                    89.5%;
                               Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     115
191
191
383
369
196
209
234
234
424
424
431
513 AA;
                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                 NCBI_TaxID-31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLG_HCVH4
Q01404;
01-JUL-1993 (
                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
SEQUENCE
                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                     RESULT 3
POLG_HCW14
ID POLG_H
AC Q01404
DT 01-JUL
DT 16-OCT
DE GENOME
DE ENVELO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                   셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CELLUIAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MARIX PROTEIN (FOTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.

MEDLINE-93019030; PubMed-1383400;

Abe K.; Inchauspe G., Fujisawa K.;

"Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.;

Gen. Virol. 73:2725-2729(1992).

-I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROFIEN BNELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                    01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66-OMB polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GF92) (GF95); Envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).
Hepatitis C virus (isolate HCV-KF) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepativirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BY THE
                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nonstructural protein.

REMOVED FROM CAPSID PROTEIN C
                                                                                                      Indels
520
56499 MW; AAl35246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC. . .)
                                                               Score 137; DB 1;
Pred. No. 4.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                              520 AA.
                                                                                                      1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BY SIMILARITY
                                                                                                                                              1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                   N-LINKED (
N-LINKED (
N-LINKED (
N-LINKED (
                                                           Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D10687; BAA01529.1; -.
                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      115
191
191
369
209
233
234
418
```

```
POLG_HCVJ7
                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                          Best
                                                                                                                                                                                                                                                                                        RESULT
     å
                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                        01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66nome polyprotein (Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2) (Fragment).
Hepatitis C virus (isolate HC-J5) (HCV).
Viruses, SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                      Ø
                                                                                                                                                                                                                                                                                                                                                                         VITOLOGY 188:331-341(1992).

-1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.
                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE-92230232; PubMed-1314459;
MEDLINE-92230232; PubMed-1314459;
MEDLINE-92230232; PubMed-1314459;
Menaka T., Fukuda S., Tsuda S.I., Mishiro S.;
"Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                 ő
                                                              Length 520;
                                                        89.5%; Score 137; DB 1; Length 52
89.3%; Pred. No. 4.4e-12;
rive 1: Mismatches 2; Indels
                                         56476 MW; 1D2BD0A6FF27349B CRC64;
                                                                                                                                                                        737 AA.
                                                                                1; Mismatches
                                                                                                   1 PKPORKTKRKAHRRPODVKFPGGGQIVG 28
                                                                                                              Query Match 89.55
Best Local Similarity 89.35
Matches 25; Conservative
                                                                                                                                                                        STANDARD;
                  NCBI_TaxID-11112;
                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                       POLG_HCVJ5
P27960;
                                                                                                                                                                                                                                                                                                                                                                  genotypes.
                              NON_TER
SEQUENCE
  CARBOHYD
            CARBOHYD
                     CARBOHYD
                                                                                                                                           SOLG HCVJ5
ID POLC
  FFFFS
                                                                                                    ð
                                                                                                                     요
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NSI and NSI (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus (isolate HC-J7) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-92230232; PubMed-1314459; Okamoto K., Lizuka H., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a heptitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                      Length 737;
                                                                                                                                                                                                                                                                                                                                                                    / Match 89.5%; Score 137; DB 1; Length 73 Local Similarity 89.3%; Pred. No. 6.3e-12; les . 25; Conservative 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                             81207 MW; 3AF699D82AD501B1 CRC64;
N-LINKED (GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                737 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                         N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL, D10077; BAA00971.1; -:
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NSI.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                             737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLG_HCVJ7
P27961;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                     CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                        Query Match
```

```
DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

15-SEP-2003 (Rel. 42, Last annotation update)

Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);

ENVELOPE 91ycoprotein E1 (GP32) (GP35); Envelope 91ycoprotein E2

(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)

EC 34.22.-); Protease/helicase NS3 (P70) (Hepacivirin)

E (EC 34.21.98); Nonstructural protein NS4A (P4); Nonstructural protein

DE NS4B (P27); Nonstructural protein NS5A (P66); Nonstructural protein

NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48);

Hepatitis C virus (isolate BK) (HCV).

C Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0
                                                                                                                                    CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ./E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSZ (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.;
Structure and organization of the hepatitis C virus genome isolated from human carriers.;
                                                                                                            REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                           POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDINE-96235224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
POlyprotein; Glycoprotein; Coat protein; Envelope protein; Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 737;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                             CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                    N-LINKED GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 137; DB 1;
Pred. No. 6.3e-12;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPORKTKRKAHRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE-91140698; PubMed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVBK
                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
SEQUENCE
                                                                                                                                                                                                         TRANSMEM
                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match
                                                                                                                                      CHAIN
                                                                                                                                                               CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 7
POLG_HCVBK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                         14441°
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
```

". NS4D.
". NS4D.
". NS4D.
". NS4D.
". PR007095; RNA\_POl\_DS\_PS.
". IPR007094; RNA\_POl\_DS\_PS.
". PF01543; HCV\_capsid; 1.
". PF01542; HCV\_core; 1.
". PF01550; HCV\_core; 1.
". PF01538; HCV\_NS1; 1.
". PF01538; HCV\_NS1; 1.
". NS02907; 1.

Pfam;

PF01006;

Pfam; Pfam;

InterPro; IPR007095; InterPro; IPR007094;

InterPro; IPR000745; InterPro; IPR001490;

InterPro; IPR002868; InterPro; IPR002166;

HCV\_capsid.

InterPro; IPR001410; InterPro; IPR002521;

MEROPS; S29.001 MEROPS; U39.001

BOHM; 1GX6;

InterPro; IPR002522; InterPro; IPR002519; InterPro; IPR002531; InterPro; IPR002518; InterPro; IPR004109;

PDB; 1JXP; 14-JAN-98 PDB; 1NS3; 08-APR-98 PDB; 1C2P; 15-NOV-00

1A1Q; 25-MAR-98

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                         X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.

X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.

X MEDLINE-98227846; PubMed-9568891;

Y An Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,

Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.,

"Complex of NS3 protease and NS4A peptide of EM strain hepatitis C

"Tomplex of NS3 protease and NS4A peptide of EM strain hepatitis C

"Tomplex of NS3 protease and NS4A peptide of EM strain hepatitis C

"Tomplex of NS3 PRAIL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE

"YDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

"NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

"C -1 CATALTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral position, Cys or Thr in P1 and Ser or Ala in P1.

"C -1 CATALLYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                             the viral
                                                                                       Love R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N., Moomaw E.W., Adachi T., Hostomska Z.; The crystal structure of hepatitis C virus NS3 proteinase reveals trypsin-like fold and a structural zinc binding site."; cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                      X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE-97015088; PubMed-8861916;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [RNA](N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A38465;
```

```
TURN
STRAND
SEQUENCE
     TURN
                            HELIX
                                                                                                                                                                                                             POLG_HCVJA
                                                                                                                                                                                                  RESULT
                                                                                                                                      ò
                                                                                                                                                          g
                                                                                                                                                                                                                                    511111
                                                                                                                      CELLULAR ANINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOR (POTENTIAL).
NONSTRUCTURAL PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS44 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF01506; HCV_NS5a; 1.
SMAR; SM0487; DEXDc; 1.
SMART: SM0487; DEXDc; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                              REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                           POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                       729
1006
1615
1862
2013
3010
                                                                                                                                                                                                                                            369
11083
11165
11165
1237
1319
1319
209
234
234
417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     645
2041
2077
2077
22529
2788
2788
1047
1063
1063
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1077
1082
1086
1090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10093
11009
11100
11120
11139
11139
11149
11168
11172
11173
                                                                                                                                                                                                                                            347
1083
1107
1165
1230
1316
196
                                                                                                                                                                                                                                                                                                                                209
234
250
305
417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2041
2240
2240
2529
2788
2788
1031
1039
1050
1059
1059
                                                                                                 3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                              ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                              TRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HELIX
STRAND
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAND
HELIX
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
STRAND
STRAND
STRAND
TURN
STRAND
STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TURN
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAND
                                                                                                                                    CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TURN
   NAME OF THE STATE OF THE STATE
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation-the European Bioinformatics institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-91088550; PubMed-2175903;
Kato N., Hilkata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
"Appanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                    89.5%; Score 137; DB 1; Length 3010; 89.3%; Pred. No. 2.8e-11; 1.ve 1; Mismatches 2; Indels
                                                                                    327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                            3010 AA
                                                                                                                                                                                      EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ.
                                                                                               Query Match
Best Local Similarity 89.55.
Conservative
                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DISCUSSION OF SEQUENCE.
                                 1202
1204
1688
1187 118
1189 119
1198 120
1203 120
1680 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11116;
                                                                                                                                                                                                                                                                                                        POLG_HCVJA
P26662;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus
```

PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28

```
ч
                                                                                                                                                 RESULT
                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Probom; PD186062; HCV_NS1; 1.

RAMART; SM00487; DEXDC; 1.

RAMART; SM00487; DEXDC; 1.

RAM Core protein; Glycoprotein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

FT INIT_MET 1 115 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPERTIDASE.

FT CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).

FT CHAIN 192 383 MAJOR ENVELOPE PROTEIN NS1 (POTENTIAL).

FT CHAIN 1007 1615 RONSTRUCTURAL PROTEIN NS2 (POTENTIAL).

FT CHAIN 1007 1615 RONSTRUCTURAL PROTEIN NS2 (POTENTIAL).

FT CHAIN 1007 1615 RONSTRUCTURAL PROTEIN NS4 (POTENTIAL).

FT CHAIN 1010 RNASTRUCTURAL PROTEIN NS4 (POTENTIAL).

FT CHAIN 3010 RNASTRUCTURAL PROTEIN NS5 (POTENTIAL).

FT CHAIN 3010 RNASTRUCTURAL PROTEIN NS5 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                . .) (POTENTIAL)
. .) (POTENTIAL)
. .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 137; DB 1; Length 3010;
Pred. No. 2.8e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                   erpro; IPR007095; RNA_pol_DS_PS.
erpro; IPR007094; RNA_pol_DS_PS.
n; PF01543; HCV_capsid; 1.
pp01543; HCV_capsid; 1.
pp01559; HCV_cenv; 1.
PF01560; HCV_NS1; 1.
PF01560; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N.LINKED
N.L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                HCV_capsid.
                                                                                                            InterPro; IPR002521; HCV_Core.
InterPro; IPR002519; HCV_Env.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR004109; HCV_NS2.
InterPro; IPR000145; HCV_NS3.
InterPro; IPR000145; HCV_NS3.
InterPro; IPR001499; HCV_NS4a.
InterPro; IPR002669; HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01206; HCV_NS5a; 1.
Pfam: PF00271; helicase_C; 1
Pfam: PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2788
v; 327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             89.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS2;
                                                                     InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
HSSP; P26663; 1JX
MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01006;
PF01001;
PF01506;
                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions and a sits content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                               01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annocation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (MS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P65); NONSTRUCTURAL (P65); NONSTRU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; Pubmed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA. PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTECTIVIT OF PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genomic structure of the human prototype strain H of hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        virus: comparison with American and Japanese isolates.".
Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                        PRT; 3011 AA
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (RNA) (N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                   POLG_HCVH
P27958;
```

ö

Gaps

ö

Indels

89.3%;

Conservative

Local Similarity les 25; Conserv

Best Loca Matches

ö Gaps ö 3033 AA

```
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              649 64
1091 109
2038 203
2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity
nes 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus.
NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLG_HCVJ8
P26661;
                                                                                                                                                                                                                                                                                                                                 CHAIN
CHAIN
CHAIN
TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 11
      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                          MEDLINE-9204440; PubMed-1658196;
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
Machida A., Miyakawa Y., Mayumi M.;
Mucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved of divergent regions.";
J Gen. Virol. 72:2697-2704(1991).
C -I- FUNCTION: THE SMALL PROPEINS NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C -I- CATALYITC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in pl and Ser or Ala in Pl'
C -I- CATALYITC ACTIVITY: N nucleoside triphosphate - N diphosphate +
          01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 21, Last sequence update)
01-40G-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP69) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
VITUSES; SERNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002166; HCV_RGRP.
IPR001650; Hellcase_C.
IPR007095; RNA_pol_DS_PS.
IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_core
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_env.
HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_capsid; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D00944; BAA00792.1; -. PIR; J01303; J01303. HSSP; P27958; 1HEI. MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002518;
                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                    NCBI_TaxID=11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -
DON WAR TO DO DO WAR TO DO WAR TO
```

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2003 (Rel. 41, Last annotations (Core protein (P22);
29-FEB-2003 (Rel. 41, Last annotations)
29-FEB-2003 (Nel); Protein P7; Nonstructural protein NS4A (P49); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48);
MS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48);
                                                                                                                                                                                                                                                                                                                                                                                                                                                     CAPSID PROTEIN C (POTEWITIAL).
MATRIX PROTEIN C (POTEWITIAL).
MAJOR ENVELOPE PROTEIN E (POTEWITIAL).
NONSTRUCTURAL PROTEIN NS1 (POTEWITIAL).
NONSTRUCTURAL PROTEIN NS2 (POTEWITIAL).
PROTEASE, HELICASE NS3 (POTEWITIAL).
NONSTRUCTURAL PROTEIN NS4A (POTEWITIAL).
NONSTRUCTURAL PROTEIN NS4A (POTEWITIAL).
RNA - DIRECTED RNA POLYMERASE (POTEWITIAL).
POTEWITIAL.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
Probom; PD186062; HCV_NS1; 1.
SWART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1 CELLULAR AMINOPEPTIDASE.
CCLLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BY SIMILARITY).
(BY SIMILARITY).
(BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CRC64:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2811 N-LINKED (ULLANDE); 329165 MW; F957F5C1A273BE9E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     89.5%; Score 137; DB 1;
89.3%; Pred. No. 2.9e-11;
11ve 1; Mismatches 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DECH BOX.
```

SIMILARITY). SIMILARITY). SIMILARITY).

(POTENTIAL). (POTENTIAL). (POTENTIAL). (POTENTIAL).

(POTENTIAL).

POTENTIAL

```
01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
61-APR-1993 (Rel. 42, Last annoctation update)
66-000 (Gontains: Capsid protein C (Core protein) (P22);
67-00 (NSI); Protein El (GP32) (GP35); Envelope glycoprotein E2
67-00 (NSI); Protein P7: Nonstructural protein NS2 (P21)
6. C3 4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein
NS48 (P77); Nonstructural protein NS58 (P56); Nonstructural protein
NS58 (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Taiwan) (HCV).
Viruses; SRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 peptide bonds in the viral
                 RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                {RNA}(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /; DB 1; Length 3vs:
. 2.9e-11;
--hea 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1A173E7E3381FD1A CRC64;
                                                              (BY
(BY
(BY
                                                           CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                           N-LINKED (GLCNAC.
N-LINKED (GL
                                                                                                                                 ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 137;
Pred. No. 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 330177 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ch 89.5%;
l Similarity 89.3%;
25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2038 203
2359 235
2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLG_HCVTW
P29846;
            CHAIN
TRANSMEM
ACT_SITE
                                                                                                      ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 12
POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@iBb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Transferase; RNA-directed RNA polymerase;
W Polyprotein; Glycoprotein; Envelope protein; Helicase, ATP-binding;
Core protein; Goat protein; Envelope protein; Helicase, ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
T INIT_MET 1 115 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
T CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).
T CHAIN 192 383 MAJOR ENVELOPE PROTEIN B; (POTENTIAL).
T CHAIN 734 1010 NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
T CHAIN 1619 NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
T CHAIN 1619 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1619 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1619 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1620 1866 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
                                                                                                                                           SEQUENCE FROM N.A.
MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
"Full-length sequence of a hepatitis C virus genome having poor
homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                  SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
Ffam; PF01642; HCV_core; 1.
Pfam; PF01642; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01001; HCV_NS2; 1.
Pfam; PF01005; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS5a; 1.
Pfam; PF01098; VIral_RGRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Interpro; IPR004109; H
Interpro; IPR000745; H
Interpro; IPR001490; H
Interpro; IPR002868; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410;
InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  interPro; IPR002166;
interPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .nterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEROPS; S29.001; -
```

OR KWW KWW CORREST OR CORREST OR

ö

Gaps

Gaps

```
between the Swiss Institute of Bioinformatics and the EMBL outstation—
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@lisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-91172826; PubMed-1848704;
MEDLINE-91172826; PubMed-1848704;
A Choo O.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,
A Gallegos C., Coil D., Medina-Selby A., Barr P.J., Weiner A.J.,
Bradley D.W., Kuo G., Houghton W.;
A Gallegos C. Coil Houghton W.;
Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1951).
Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1951).
C-1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
C-1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
C-1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
6enome polyprotein (Contains: Capsid protein C (Core protein) (P22);
6enome polyprotein El (GP32) (GP35); Envelope glycoprotein E2
6press (GP68) (GR10) (NS1); Protein P1 Nonstructural protein NS4 (P4); Nonstructural protein NS4 (P4); Nonstructural protein NS4 (P4); Nonstructural protein NS5 (P56); Nonstructural protein NS5 (P56); Nonstructural protein NS5 (P66) (P77); Nonstructural protein NS5 (P66); NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (RNA)(N).

SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                   (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                       (POTENTIAL
                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                 Length 3010;
     N-LINKED (GLCNAC. ..) (POTENN N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                         Score 130; DB 1; L. Pred. No. 2.8e-10; 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2 KPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   623
645
2041
2077
2240
2788
1; 327047
                                                                                                                                                                                                                                                                                                                                                                       sh 85.0%;
l Similarity 88.9%;
24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11104;
                                                                              540
556
576
623
645
2041
2047
2240
22529
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLG_HCV1
P26664;
                                                                                                                                                CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local (
                                                          CARBOHYD
                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCV1
          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE.

CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROYEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
  E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                          SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL,
CHARGE RELAY SYSTEM (
CHARGE RELAY SYSTEM (
CHARGE RELAY SYSTEM (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                      EMBL; M84754; -; NOT_ANNOTATED_CDS
PIR; A40244; GNWVTW.
PDB; 1N64; 25-FEB-03.
PDB; 1NS3; 08-APR-98.
PROTEIN M AND GLYCOPROTEIN PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             110863
2014
2014
1083
111007
11230
1233
203
4117
4117
4117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ilensib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDILINE-92113549; PubMed-1662697;
MEDILINE-92113549; PubMed-1662697;
"Comparative sequence analysis of the long repeat regions and adjoining parts of the long unique regions in the genomes of herpes simplex viruses types 1 and 2."
J. Gen. Virol. 72:3057-3075(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL: D10471; BAA01269.1; -.
EMBL: Z86099; CAB06702.1; -.
PIR: JQ1498; WMBEXA.
EALLy protein; Transcription regulation; Activator; DNA-binding.
SEQUENCE 512 AA; 54958 MW; 459651470A503BA7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FURI_HUMAN STANDARD; PRT; 794 AA.
P09958; Q14336;
01-MAT-1989 (Rel. 10, Created)
01-MAR-1989 (Rel. 14, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Furin precursor (EC 3.4.21.75) (Paired basic amino acid residue
                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ..
                                          Score 129; DB 1; Length 3011;
Pred, No. 4e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 53; DB 1; Length 512;
Pred. No. 4.6;
3; Mismatches 13; Indels
        327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                 Herpes simplex virus (type 2 / strain HG52).
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
NCBI_TaxID=10315;
                                                                                                                                                                                                                                                                                                            01-DEC-1992 (Rel. 24, Created)
01-DEC-1992 (Rel. 24, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Transcriptional regulator IE63 (VMW63) (ICP27).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      160 PKPRRRVSRNAHNQGGRHPASARTDGPGATHG 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  58
                                                                                                                                                                                                                                                                            512 AA
                                                                                       2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTKRKAH-----RRPQDVKFPGGGQIVG
                                                                                                                                                   5 PKPQKKNKRNTNRRPQDVKFPGGGIVG 32
                                                                                                                             1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                            Match 84.3%;
Local Similarity 82.1%;
les 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 37.5%;
Matches 12; Conservative
                                                                                                                                                                                                                                                                          STANDARD;
          3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                      IE63_HSV2H
P28276:
        SEQUENCE
                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 15
FURI_HUMAN
                                                                                       Matches
                                                                        Best
                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                    윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENCYCLIN (POTENTIAL).
NAJOR ENCYCLIN (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REMOVED FROM CAPSID PROTEIN C BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
CHARGE RELAY SYSTEM ()
CHARGE RELAY SYSTEM ()
CHARGE RELAY SYSTEM ()
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                       InterPro; IPR002868; HCV_NS5a.
InterPro; IPR002166; HCV_RdRP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
                                                                                                                           HCV_capsid.
                                                                                                                                                                                                                                                   HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                     1. PPO1543; HCV_capsid; 1. PPO1542; HCV_capsid; 1. PPO1542; HCV_cace; 1. PPO1559; HCV_cac; 1. PPO1560; HCV_NS1; 1. PPO1509; HCV_NS2; 1. PPO1000; HCV_NS4; 1. PPO1000; HCV_NS4s; 1. PPO1001; HCV_NS1s; 
                                                                                                                                    InterPro; INTERPOSS 1; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      645
2041
2077
2240
2364
2789
                                                                                                                                                                                                                           InterPro; IPR004109; H
InterPro; IPR000745; H
InterPro; IPR001490; H
InterPro; IPR002868; H
PIR; A39166; GNWVC3.
PDB; 1A1V; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
                                                                                                    InterPro; IPR001410;
InterPro; IPR002522;
                                                           MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 T_SITE_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ProDom;
SMART; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                             etam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              fam;
```

ä

between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; X17094; CAA34948.1; -. EMBL; X04329; CAA27860.1; -. EMBL; A06939; CAA00605.1; -.

Genew; HGNC:8568; FURIN.

MIM; 136950;

PIR; A39552; KXHUF. HSSP; Q99405; 1MPT. MEROPS; S08.071; -.

```
Leduc R., Wolloy S.S., Thoras G.;
Leduc R., Wolloy S.S., Thoras Leduc R., Molloy S.S., Thoras G.;
Activation of human furin precursor processing endoprotease occurs by an intramolecular autoproteolytic cleavage.";
J. Biol. Chem. 267:14304-14308 (1992).

I. FUNCTION: FURIN IS LIKELY TO REPRESENT THE UBIOUTOUS ENDOPROTEASE ACTIVITY WITHIN CONSTITUTIVE SECRETORY PATHWAYS AND CAPABLE OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF.

I. CATALYTIC ACTIVITY: Release of mature proteins from their proproteins by cleavage of Arg Naa-Yaa-Arg-1-Zaa bonds, where Xaa can be any amino acid and Yaa is Arg or Lys. Releases albumin, complement C3 and von Willebrand factor from their
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Roebroek A.J.M., Schalken J.A., Leunissen J.A.M., Onnekink C., Bloemers H.P.J., van de Ven W.J.M.; "Evolutionary conserved close linkage of the c-fes/fps proto-oncogene and genetic sequences encoding a receptor-like protein."; EMBO J. 5:2197-2202(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3D-STRUCTURE MODELING OF CATALYTIC DOMAIN.
MEDLINE-94291619; PubMed-8020465;
Siezen R.J., Creemers J.W.M., van de Ven W.J.M.;
Homology modelling of the catalytic domain of human furin. A model for the eukaryotic subtilisin-like proprotein convertases.";
Eur. J. Blochem. 222:255-266(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                 "CDNA and gene structure for a human subtilisin-like protease with cleavage specificity for paired basic amino acid residues."; DNA Cell Biol. 10:319-328(1991).
                                                               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBI_TaxID=9606;
                                                                                                                                                                                                                              van den Ouweland A.M.W., van Duijnhoven H.L.P., Keizer G.D., Dorssers L.C.J., van de Ven W.J.M.; "Structural homology between the human fur gene product and the subtilisin-like protease encoded by yeast KEX2."; Nucleic Acids Res. 18:664-664(1990).
                                                                                                                                                                                                                                                                                                                                                                                                        Barr P.J., Mason O.B., Landsberg K.E., Wong P.A., Kiefer M.C.,
enzyme) (PACE) (Dibasic processing enzyme).
PACE OR FUR.
                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-91321735; PubMed-1713771;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 296-794 FROM N.A. MEDLINE-87053858; PubMed=3023061;
                                                                                                                                                                                                       MEDLINE-90175002; Pubmed-2408021;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-92332543; PubMed=1629222;
                                                      Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                       TISSUE-Blood;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROCESSING
           cleaving FURIN OR
                                                                                                                                                                                                                                                                                                                                                                                                                                 Brake A.
```

PROSITE; PS00138; SUBTILASE\_SER; 1. Hydrolase; Serine protease; Transmembrane; Glycoprotein; Signal;

ProDom; PDUUU...;
SWART; SW00261; FU; 2.
PROSITE; PS00136; SUBTILASE\_ASP; 1.
PROSITE; PS00137; SUBTILASE\_RIS; 1.
PROSITE; PS00137; SUBTILASE\_ERR; 1.

POTENTIAL. FURIN. CYS-RICH. POTENTIAL

Zymogen; Calcium

SIGNAL

GO: GO:0005794; C:Golgi apparatus; TAS.
GO: GO:0004276; F:furin activity; TAS.
GO: GO:0006508; F:furin activity; TAS.
GO: GO:0006508; P:protecl-cell signaling; TAS.
GO: GO:0006508; P:proteclysis and peptidolysis; TAS.
InterPro: IPR005212; Furin repeat.
InterPro: IPR00209; Peptidase\_S8.
InterPro: IPR00209; Peptidase\_S8.
InterPro: IPR00209; Peptidase\_S8.
InterPro: IPR007213; Purpprotein; PARTIAL.
FRINTS: PR00723; SUBTILISIN.
FRIDOM: PD00773; P\_Gomain; 1.

(BY SIMILARITY)
(BY SIMILARITY)
(BY SIMILARITY)

CHARGE RELAY SYSTEM CHARGE RELAY SYSTEM CHARGE RELAY SYSTEM

24 107 705 705 705 1153 1153 1194 360 380 387 440

ACT\_SITE ACT\_SITE DISULFID DISULFID CARBOHYD CARBOHYD

TRANSMEM ACT\_SITE

DOMAIN PROPEP

```
ö
                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
CLEAVAGE (SECOND AUTO-).
CLEAVAGE (FIRST AUTO-).
CELA ATTACHMENT SITE (POTENTIAL).
CELL SURFACE SIGNAL.
TRANS GOLGI NETWORK SIGNAL.
TRANS GOLGI NETWORK SIGNAL.
                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                      Score 53; DB 1; Length 794; Pred. No. 7.4;
                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                  5,
                                                                                                                                                                                                                                                                                                                                                  Mismatches
                                                                                                                                                                                                                                                                                                                                                                                               7, 2003, 11:20:06
                                                                                                                                                                                                                                                                                                                                                                      :|:||| :: | | |||
102 KRRTKRDVYQEPTDPKFP 119
                                                                                                                                                                                                                                                                                                                                                               21
                                                                                                                                                                                                                                                                                                                          W.
                                                                                                                                                                                                                                                                                                                                      34.6%;
50.0%;
                                                                                                                                                                                                                                                                                                                                                               4 QRKTKRKAHRRPQDVKFP
                                                                                                                                                                                                                                                                                                                         86678
                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                  553
75
107
500
762
7762
                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
9; Conserve
                                                                                                                                                                                                                                                                                                                                                                                              Search completed: August
Job time: 5.90909 secs
                                                                                                                                                                                                                                                                                                                         794 AA;
                                                                                                                                                                                                                                                                                                                         SEQUENCE
QQ
                                                                                                                                                                                                                                                                                                                                                               ö
```

COFACTOR: CALCION-DEPENDENT.
ENZYME REGULATION: COULD BE INHIBITED BY THE NOT SECONDLY CLEAVED

respective precursors.

PROPEPTIDE

-

++ -

,99999999999

-!- SUBCELLULAR LOCATION: SEEMS TO BE LOCALIZED INTRACELLULARLY TO THE TRANS GOLGI NETWORK. PROPEPTIDE CLEAVAGE IS A PREREQUISITE FOR EXIT OF FURIN MOLECULES OUT OF THE ENDOPLASMIC RETICULM (ER). SECOND CLEAVAGE IN THE PROPEPTIDE OCCUR IN THE TRANS GOLGI NETWORK (TGN), WHICH IS FOLLOWED BY THE RELEASE OF THE PROPEPTIDE BOUND TO FURIN AND THE ACTIVATION OF FURIN.

-!- TISSUE SPECIFICITY: SEEMS TO BE EXPRESSED UBIQUITOUSLY.

-!- DOMAIN: CONTAINS A CYTOPLASMIC DOMAIN RESPONSIBLE FOR ITS TGN LOCALIZATION AND RECYCLING FROM THE CELL SURFACE.

-!- PTW: THE PROPEPTIDE IS AUTOCATALYTICALLY REMOVED THROUGH AN INTRAMOLECULAR CLEAVAGE PROBABLY IN THE ENDOPLASMIC RETICULUM (ER). IN THE TGN THE SECOND CLEAVAGE IN THE PROPEPTIDE COULD LEAD -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY -!- SIMILARITY: Contains 1 homo B/P domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration

```
August 7, 2003, 11:05:41; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                    tal number of hits satisfying chosen parameters:
                                                                                                                                                                  US-09-491-146A-33
153
1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                      830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                               OM protein - protein search, using sw model
                                                                                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SPTREMBL_23:*

1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_plant: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sb_mhc:*
                                                                                                                                                                  Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Database :
                                                                                                                                                                                                                                                                                      Searched:
                                                                                                  Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		Q9pxn3 hepatitis c	068575 hepatitis c	_	_	Q81541 hepatitis c	068311 hepatitis c	Q68306 hepatitis c	Q68309 hepatitis c		Q8jyr8 hepatitis c	Q8jyr5 hepatitis c	Q8jyr7 hepatitis c	Q8jyr6 hepatitis c	OFR712 henatitie c
P90364	081831	Q9PXN3	968575	Q91AU0	081989	081541	068311	Ø68306	068309	Q8JYR9	Q8JYR8	Q8JYR5	Q8JYR7	QBJYR6	068712
12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
105	106	125	191	3010	3010	3010	45	45	46	09	61	61	61	62	74
90.2	90.5	90.5	90.2	90.2	90.3	90.5	89.5	89.5	89.5	89.5	89.5	89.5	89:2	89.5	89.5
138	138	138	138	138	138	138	137	137	137	137	137	137	. 137	137	137
	8	٠ ،	4	S	9	7	80	6	10	11	12	13	14	15	16
	90.2 105 12 p90364	90.2 105 12 P90364 P90364 P90364 90.2 106 12 Q81831	90.2 105 12 P90364 P90364 90.2 106 12 QB1831 QB1831 90.2 125 12 Q9PKN3 Q9PKN3	90.2 105 12 P90364 P90364 P90364 90.2 106 12 QB1831 QB1831 QB1831 QB7813 P90.2 125 12 Q9PXN3 Q9PXN3 Q9PXN3 Q8F873 Q8F873 Q8F878	90.2 105 12 P90364 P90364 90.2 106 12 QBH831 QBH831 90.2 125 12 QSPXN3 QSPXN3 90.2 191 12 QGBS75 QGBS75 90.2 3010 12 Q91AU0 Q91AU0	90.2 105 12 P90364 P90364 90.2 106 12 081831 90.2 125 12 Q9PXN3 Q9PXN3 90.2 191 12 Q9575 G68575 90.2 3010 12 Q91AU0 Q919aU0 90.2 3010 12 Q81989	90.2 105 12 P90364 P90364 90.2 106 12 OH1831 QPRN3 90.2 125 12 QPRN3 90.2 191 12 Q68575 Q68575 90.2 3010 12 Q91A00 Q91Au0 90.2 3010 12 Q81989 Q81989 90.2 3010 12 Q81541 Q81541	90.2 105 12 P90364 90.2 106 12 QBH831 QBH831 90.2 125 12 QSPXN3 QSPXN3 90.2 191 12 Q68575 Q68575 90.2 3010 12 Q91,000 Q91,00 90.2 3010 12 Q81,989 Q81541 89.5 45 12 Q81311 Q6831541	90.2 105 12 P90364 P90364 90.2 106 12 QBH831 QBH831 90.2 125 12 QBFXN3 Q68575 90.2 3010 12 Q91A00 Q81989 90.2 3010 12 Q81989 Q81989 90.2 3010 12 Q81989 Q81541 89.5 45 12 Q68311 Q68311	90.2 105 12 P90364 P90364 90.2 106 12 Q9L831 Q9PxN3 90.2 125 12 Q9PxN3 Q9PxN3 90.2 3010 12 Q9L800 Q91380 90.2 3010 12 Q81389 Q81389 90.2 3010 12 Q81389 Q81581 89.5 45 12 Q68311 Q68306 89.5 46 12 Q68309	90.2 105 12 P90364 90.2 106 12 QBH831 QBH831 90.2 125 12 Q9PKN3 Q9PKN3 90.2 3010 12 Q9LA00 90.2 3010 12 Q9LA00 90.2 3010 12 Q9LA00 90.2 3010 12 Q81989 Q81541 89.5 45 12 Q68311 Q68306 89.5 46 12 Q68306 Q68306 89.5 60 12 Q817K9	90.2 105 12 P90364 90.2 106 12 QBH831 90.2 125 12 QBFXN3 90.2 1311 12 QBB575 90.2 3010 12 QB1989 90.3 3010 12 QB1989 90.5 45 12 QB8309 90.68309 90.68309 90.68309 90.68309 90.7 QB19789 90.8 QB19789	90.2 105 12 P90364 90.2 106 12 OBH831 90.2 125 12 OPENN3 90.2 191 12 OBB575 90.2 3010 12 OBH870 90.2 3010 12 OBH89 90.2 3010 12 OBH89 90.3 3010 12 OBH89 90.5 45 12 OBH89 90.5 46 12 OBH88 90.6 301 90.7 301 90.8 301	90.2 105 12 P90364 90.2 106 12 OBH831 90.2 125 12 OSPXN3 90.2 191 12 OSB575 90.2 3010 12 OSH599 90.2 3010 12 OSH599 90.5 45 12 OSB306 89.5 45 12 OSB306 89.5 60 12 OSB306 89.5 60 12 OSB3789 89.5 61 12 OSB3789	105 12 P90364 P90364 106 12 OB1831 OB

hepatitis		hepatitis	hepatitis	hepatit	hepatitis	hepat	hepatitis	hepatitis	hepatitis	hepatitis	hepatit	hepatitis	Q8qp54 hepatitis c	hepatitis	hepatitis	hepatitis		hepatitis	hepatit	Q8qp55 hepatitis c			hepatitis		hepatitis	hepatit	hepatitis	tis	
6868	068708	6867	6868	6870	6868	6870	6868	8JYS	QBJYS1	8126	Q9QT61	8QP7	8QP5	90T5	ворь	8QP8	9QTS	вор9	8QP6	8QP5	090P60	8QP5	Q9QT58	90Б6	Q8QP61	Q9QT52	09QT54	Q8QP59	ALIGNMENTS
4	74 12	7	7	4	4	4	1	9	1	83 1		00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	
σ.	89.5	φ.	6	Ġ.	φ.	6	σ.	6	6	6	6	6	6	6	6	Ġ.	ó.	6	ę.	6	ď.	σ,	6	6	6	σ.	σ.	φ.	
3	137	'n	m	3	3	m	m	m	m	m	e	ന	m	S	m	က	3	m	m	m	'n	3	m	m	m	3	'n	n	
17	18	19	20	. 21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	ъ 6	40	41	42	43	44	45	

	ARY; PRT; 105 AA.	03,	03,	el. 22, Last annotation update)	Core protein (Genome polyprotein) (Fragment).		Viruses; sakna positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus					"Determination of nine genotypes of hepatitis C virus using PCR		Thesis (1995), Nagoya City University Medical School.			MEDLINE=96305314; Pubmed=8712927;	Ohno T., Mizokami M., Saleh M.G., Orito E., Ohba KI., Wu RR.,	Kolde T., Tibbs C.J., Nouri-Aria K.T., Tokudome S., Williams R.;	"Usefulness and limitation of phylogenetic analysis for hepatitis C	virus core region: application to isolates from Egiptian and Yemeni		11-1113(1996).	-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A	LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:	PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF	PROTEIN C AND MRNA (BY SIMILARITY).	19.1;	HCV_capsid.	apsid; 1.				11870 MW; C7BA40B284025A49 CRC64;
	PRELIMINARY;			(TrEMBLrel.	(Genome pol)	virus.	NA positive-	1103:		M N.A.	okami M.;	on of nine go		), Nagoya Ci		M N.A.	5314; PubMed	okami M., Sa	bbs C.J., No	and limitatio	egion: appli		141:1101-11	THE VIRION (	EIN ENVELOPE	M AND GLYCOP!	C AND MRNA ()	EMBL; D82034; BAA11519.1;	InterPro; IPR002522; HCV_capsid.	Pfam; PF01543; HCV_caps1d;		1 1	105 105	
LT 1 64	P90364 P90364:	01-MAY-1997		01-OCT-2002	Core protein	Hepatitis C virus.	Viruses; ssk	NCBI TaxID-11103:	[1]	SEQUENCE FROM N.A.	Ohno T., Mizokami M.;	"Determinati	method.";	Thesis (1995	[3]	SEQUENCE FROM N.A.	MEDLINE-9630	Ohno T., Miz	Koide T., Ti	"Usefulness	virus core r	patients.";	Arch. Virol.	-1- SUBUNIT:	LIPOPROT	PROTEIN	PROTEIN	EMBL; D82034	InterPro; IP	Pfam; PF0154	Polyprotein.	NON_TER		SEQUENCE
RESULT P90364	A D	DT	ΔŢ	TO.	DE	S 6	3 5	88	RN	ЯР	RA	R	RT	RL	RN	RP	RX	RA A	RA	RŢ	RŢ	RŢ	RL	ပ္ပ	႘	ပ္ပ	ပ္ပ	DR	DR	DR	ΚW	FT	F. C	S.

ð

SCORPTPROSCO

```
Length 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      191
20840 MW; 4AAE63444D8329E2 CRC64;
                                                                                                                                                                                                                                                                                                                                                         Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Created)
Last sequence update)
Last annotation update)
                                                                                3
                               Score 138; DB 12;
Pred. No. 1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 90.2%; Score 138; DB 12; Best Local Similarity 89.3%; Pred. No. 1.6e-12; Matches 25; Conservative 1; Mismatches 2;
                                                      le-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                 Core protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                             191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 3010 AA
                                                                                   1; Mismatches
                                                                                                                                 1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                           5 PKPQRKTKRNTYRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROTEIN M AND GIXCOPROTEIN E. THE N
PROTEIN C AND MRNA (BY SIMILARITY).
EMB.; U55284; AABOO215.1; -.
InterPro; IPRO02522; HCV_capsid.
InterPro; IPRO02521; HCV_core.
Pfam; PF01543; HCV_copeid; 1.
                            90.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23, 1
                                                                                                                                                                                                                                                                                                                                                           (TremBLrel. 01, (TremBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                      01-NOV-1996 (TYEMBLEEL. 01,
01-NOV-1996 (TYEMBLEEL. 01,
01-OCT-2002 (TYEMBLEEL 22,
                                                         Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE 191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    091AU0;
01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus
                                    Query Match
                                                                                                                                                                                                                                                                                                             068575
                                                                                                                                                                                                                                                                                                                                       068575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       091AU0
                                                                                                                                                                                                                                                                                                                  SO THE PROPERTY OF THE PROPERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEGOAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEGCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL: D49465; BAA08439.1; --
InterPro; IPR002522; HCV_capsid.
Pfam: PF01543; HCV_capsid.
                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE-94351179; PubMed-7520922;
MAEDLINE-94351179; PubMed-7520922;
NARAZAWA T., RATO N., Ohkoshi S., Shibuya A., Shimotohno K.;
J. Hepatol. 20:623-629(1944).
I. Hepatol. 20:623-629(1944).
I. POBROTEIN THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MANA (BY SIMILARITY).
InterPro; IPR002522; HCV_capsid,
PP01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
E2/NS1 protein (Genome polyprotein) (Fragment).
Hepatitis C virus type 2.
Hepatitis C virus type 2.
Hepacitis SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ohno T., Mizokami M.; **Determination of nine genotypes of hepatitis C virus using PCR ***Determination of nine genotypes of hepatitis C virus using PCR
                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 138; DB 12; Length 106; Pred. No. 8.8e-13;
     Length 105;
                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
(Isolate USA8) genomic RNA (Genome polyprotein) (Fragment).
                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             106 106
106 AA; 12001 MW; 25D0D5414B3EA9DC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     125 AA; 14220 MW; D945CB60661797E3 CRC64;
Score 138; DB 12;
Pred. No. 8.7e-13;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               125 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. 8.86
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                 1 PKPORKTKRKAHRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                    Created)
                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
     90.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                            01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-OCT-2002 (TrEMBLrel. 22,
  Ouery Match
Best Local Similarity 89.33
Matches 25; Conservative
                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-40271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein.
SEQUENCE 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-USAB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                              081831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9PXN3
                                                                                                                                                                                                                               RESULT 2
Q81831
```

ò

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-HCV-SI,
MEDLINE-21440119; PubMed-11556407;
Lim S.-P., Khu Y.L., Hong W.J., Tay A., Ting A.E., Lim S.G., Tan Y.H.;
Lim S.-P., Khu W.L., Hong W.J., Tay A., Ting A.E., Lim S.G., Tan Y.H.;
"Identification and molecular characterization of the complete genome
of a Singapore isolate of hepatitis C virus: sequence comparison with
other strains and phylogenetic analysis.";
Virus Genes 23:89-95(2001).
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 191;
Indels
```

```
single Japanese carrier in Nagasaki prefecture and genome analysis of E1 and E2/NS1 envelope glycoprotein regions.";
Jpn. J. Trop. Med. Hyg. 22:169-177(1994).
-1- SUBUNIT: THE VIRION OF THIS VIRIOS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL: D63857; BAA09919-1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J., Nakazawa T., Hijikata M., Ishimura Y., Shimotohno K.; "Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among infected individuals.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             12; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hydrolase, Nonstructural protein, Polyprotein, RNA-directed RNA polymerase, Transferase, Transmembrane. SEQUENCE 3010 AA, 327503 MW, C7BDB38169D6E3CF CRC64,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 138; DB 12
Pred. No. 3e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequ
01-MAR-2003 (TrEMBLrel. 23, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001650; Helicase_C.
InterPro; IPR0070095; RNA_POL_DS_PS.
InterPro; IPR007004; RNA_POl_PSvir.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-JT';
MEDLINE-92295714; PubMed-1318627;
                                                                                                                                                                                                                                                                                                                                                                                        HCV_capsid.
HCV_core.
HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01560; HVV_NS1; 1. Pfam; PF01560; HVV_NS1; 1. Pfam; PF01289; HVV_NS2; 1. Pfam; PF01007; HVV_NS3; 1. Pfam; PF01001; HVV_NS4s; 1. Pfam; PF010101; HVV_NS5s; 1. Pfam; PF00271; helicase_C; 1. Pfam; PF00998; Viral_RRRP; 1. Probom; PD186062; HCV_NS1; 1. Probom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS4b.
HCV_NS5a.
HCV_RdRP.
Helicase_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         90.2%;
ilarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002521;
InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR004109;
InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002868
InterPro; IPR002166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
nes 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q81541
Q81541;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          081541
        SO WEER RANGED BY REAL PROPERTY OF THE SO OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HER REAR TO A REAR THE A REAR THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Coat protein; Envelope protein; Glycoprotein; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane, SEQUENCE 3010 AA; 326793 MW; 3D89304314F9F795 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                        ALTEE (MAR-2001) to the EMBL/GenBank/DDBJ databases.
SUBUNIT: THE VURLON OF THIS VIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-HCV-N; Zheng W.Z.; "Genotype identification of hepatitis c virus (HCV) isolated from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
El and E2/NS1 envelope glycoprotein (Genome polyprotein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zheng W.-Y.Z.;
Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 138; DB 12
Pred. No. 3e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
||||||||||
5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                    EMBL; AF356827; AAL00900.1; -.
InterPro; IPR000345; CytC_heme_bind.
InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROSITE; PS00190; CYTOCHROME_C; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR007095; RNA_pol_DS_PS. IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002522; HCV_capsid. IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002519; HCV_env.
IPR002518; HCV_NS1.
IPR002518; HCV_NS2.
IPR004109; HCV_NS3.
IPR000145; HCV_NS4a.
IPR001490; HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002166; HCV_RdRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 90.2%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; V1ral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PF01543; HCV_caps1d;
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 89.3 es 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF01001; HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-HCV-N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                            Lim S.P.;
Submitted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             nterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q81989
Q81989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 6
081989
AC 0819
AC 0819
DT 01-N
DT 01-N
DT 01-N
DT 01-N
DT 01-N
DE 01-
```

4

ö

Gaps

ö

```
STRAIN-HCV-BB5;
Songsivilai S., Kanistanon D., Kunkitti R.;
Identification and characterisation of Thai isolates of hepatitis C
       Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GIVCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; U23749; AAA565051;
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         virus.";
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
--- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GINCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment)
Hepatitis C virus.
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                             Score 137; DB 12; Length 45;
Pred. No. 5e-13;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                   CC527167096AAA81 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CC4C21ED236AAA81 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            068306 PRELIMINARY; PRT; 45 AA. 068306; 0.001-NOV-1996 (TrEMBLrel. 01, Created) 01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) 02-0cr protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 137; DB 12
Pred. No. 5e-13;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  46 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL, U23744; AAA65051.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                            45
5015 MW;
                                                                                                                                                                                                                                                                                                                89.5%;
ilarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       45 45
45 AA; 5030 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                         45
45 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                               Polyprotein.
NON_TER
SEQUENCE 45
                                                                                                                                                                                                                                                                                                                Query Match
Best Local Simi
Matches 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q68309;
Q68309;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 10
068309
1D 06830
AC 06830
DT 01-NO
DT 01-NO
DT 01-NO
DT 01-NO
DE COPE
OS VITUS
OC HEPAC
OC HEPAC
OX HEPAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 9
Q68306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
             NO CCC CCC REAL RESERVENCE CCC CCC REAL REAL RESERVENCE CCC CCC REAL RESERVENCE CCC RE
                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ID DATE OF THE STANDARD OF THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                    4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               polymerase; Transferase; Transmembrane
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ပ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROSITE; PSS0507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                     -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
STRAIN-HCV-BB37;
Songsivilai S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       05F0B2102CF9DD9D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No. 3e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1; Mismatches
                                                                                                                                                                                                                                                         45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROTEIN M AND GLYCOPROTEIN E. THE NI
PROTEIN C AND MRNA (BY SIMILARITY).
J. D11355; BAA18894.1; -...
P. P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NS4B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last seq
01-OCT-2002 (TrEMBLrel. 22, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                    InterPro; IPR002522; HCV_capsid
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RNA-directed RNA
                                                                                                                                                                                                                                                    IPR002519; HCV_env.
IPR002531; HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   90.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF01560; HCV_NS1; 1.
PF01538; HCV_NS2; 1.
PF02907; HCV_NS3; 1.
PF01006; HCV_NS4a; 1.
PF01001; HCV_NS4b; 1.
PF01506; HCV_NS5a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       326564
                                                                                                                                                                          InterPro; IPR001410; DEAD.
Virus Res. 23:39-53(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 89.3 (atches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1006
1615
1862
2013
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein;
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               068311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
```

ö

Gaps

ö

Length 45; Indels

4

```
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                      InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AF506620; AAM33396.1; -
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genome polyprotein (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                89.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
6897 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (TrEMBLrel. 22, (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 89.5
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 89.3
Les 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-MAR-2003 (TrEMBLrel.
                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 AA;
                                          NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-RIG290;
                                                                                                                                                                                                                                                                                                                                Polyprotein.
NON_TER
SEQUENCE 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-0CT-2002
01-0CT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q8JYR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q8JYR7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q8JYR7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q8JYR.
        STEE BEAUTION OF STEEL SOLUTION OF STEEL SOLUTIO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
SEQUENCE FROM N.A.
STRAIN-HCV BB10;
SORGSIVILAI S., Kanistanon D., Kunkitti R.;
Identification and characterisation of Thai isolates of hepatitis C
                                                                                                                          K
                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                              Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GIVCORROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MANA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Shustov A.V., Gavillova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/Genbank/DbbJ databases.
-1- Subunit THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                   Length 46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 137; DB 12; Length 60;
Pred. No. 6.8e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                       ABDC4C21ED236AAA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B5691CBE8F26F2F1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                             Score 137; DB 12;
Pred. No. 5.1e-13;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          60 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. 6.80
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AF506616; AAM33392.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                    InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Genome polyprotein (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              60
6686 MW;
                                                                                                                                                                                                                                                                                      46
5129 MW;
                                                                                                                                                                                                                                                                                                                                             / Match 89.5%;
Local Similarity 89.3%;
les 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       89.5%;
llarity 89.3%;
Conservative
                                                                                                                                                                                                 EMBL; U23747; AAA65054.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                       46 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 60 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
es 25, Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-RIG306;
                                                                                                                                                                                                                                                               Polyprotein.
NON_TER 46
SEQUENCE 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q8JYR9
Q8JYR9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QBJYR8
QBJYR8;
                                                                              virus.
                                                                                                                                                                                                                                                                                                                                                                    Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 11
Q8JYR9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08JYR8
ID 08
AC 08
DT 01
DT 01
DE Ge
  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databbases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                  QF
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                        Score 137; DB 12; Length 61;
Pred. No. 6.9e-13;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                    61 61
61 AA; 6897 MW; EC656DC79E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EC656DC79E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (TrEMBLrel. 22, Created)
(TrEMBLrel. 22, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 137; DB 12;
Pred. No. 6.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Mismatches
                                                                                                                                                                                                                                                                                                                                                    28
                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last seq
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                               Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitteed (Apr-2002) to the EMBL/GenBank/DDBs databases.
-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).

InterPro; IPR002522; HCV_capsid.
Pfam: PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                           Viruses; ssrNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genome polyprotein (Fragment).
Heparitis C virus.
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                   Score 137; DB 12; Length 61;
Pred. No. 6.9e-13;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 89.5%; Score 137; DB 12; Length 62; Best Local Similarity 89.3%; Pred. No. 7e-13; Matches 25; Conservative 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                 61 61 .
61 AA; 6869 WW; EC657F706E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62 62
62 AA; 7053 MW; EEEC656DC79E8F26 CRC64;
 01-MAR-2003 (TrEMBLrel. 23, Last annotation update) Genome polyprotein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-0CT-2002 (TrEMBLrel. 22, Created)
01-0CT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL, AF506619; AAM33395.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                     [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                    Hepatitis C virus
                                                                                  NCBI_TaxID-11103;
                                                                                                                                   STRAIN-RIG288;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-RIG289;
                                                                                                                                                                                                                                                                                                                Polyprotein.
NON_TER
SEQUENCE 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
NON_TER
SEQUENCE 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OBJYR6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 15
õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

Search completed: August 7, 2003, 11:19:03 Job time : 25.6364 secs

us-09-491-146a-33.rai

```
Sequence 33, Appl Sequence 28, Appl Sequence 191, Appl Sequence 192, Appl Sequence 36, Appl Sequence 2, Appl Sequence 2, Appl Sequence 3, Appl Sequence 2, Appl Sequence 3, Appl Sequence 10, Appl Sequence 14, Appl Sequence 8, Ap
                                                                                                                                                                                                                                                 August 7, 2003, 11:07:41; Search time 10.5455 Seconds (without alignments) 112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2, 1
46,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1: /cgn2_6/ptodata/1/laa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/laa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/laa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/laa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/laa/pcTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/laa/pcTUS_COMB.pep:*
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-921-887-33
US-08-921-887-28
US-08-921-887-28
US-08-635-886C-191
US-08-380-160-6
US-09-380-160-2
US-09-380-160-2
US-09-380-160-1
US-08-380-160-1
US-08-380-160-1
US-08-380-160-1
US-08-08-050-3
US-08-050-050-3
US-08-050-050-3
US-08-262-037-26
PCT-US93-080-88-23
PCT-US93-080-88-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-836-075A-10
US-08-635-886C-198
US-08-635-886C-232
US-08-835-886C-232
US-08-324-977-8
US-08-384-616-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-501-195-2
US-08-836-075A-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUMMARIES
                                                                                                                                                                                - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-491-146A-33
153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Scoring table:
                                                                                                                                                                            OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Searched:
                                                                                                                                                                                                                                                                    Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ş
```

Applications of Applications o	NG THE SAM	<b>.</b>
60, 12, 15, 15, 15, 16, 16, 16,	MAKI	Gaps
Sednence Sed	FOR	ő
8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	METHOD	Length 28, Indels
9.5 138 3 US-08-836-075A-60 19.5 140 2 US-08-501-195-1 19.5 154 3 US-08-501-195-1 19.5 154 3 US-08-51-2 19.5 194 0 US-08-290-665A-155 19.5 191 2 US-08-290-665A-155 19.5 191 2 US-08-290-665A-156 19.5 191 2 US-08-290-665A-156 19.5 191 2 US-08-290-665A-161 19.5 191 2 US-08-290-665A-163 19.5 191 2 US-08-290-665A-165 19.5 191 2 US-08-290-665A-165 19.5 191 2 US-08-290-665A-165 19.5 191 2 US-08-290-665A-165 19.5 191 2 US-08-290-665A-165	NO SO A CONTROL OF STANDARY OF SO A CONTROL OF STANDARY OF STANDAR	internal :: :patitis virus 100.0%; Score 153; DB 3; L ity 100.0%; Pred. No. 5.8e-15; :servative 0; Mismatches 0;
	-33 (APD11C; APD11C; APD11C; APD11C; FORMATIOI TI FEELI TINVENTIC FORMATIOI TINVENTIC FORMATIOI TINVENTIC FORMATIOI TINVENTIC FORMATIOI TO T	306
	UGENCE 33 APP11ca -921-887-33 UGENCE 33 APP11ca ent NO. 66307701 APPLICANT: KHUDY APPLICANT: KHUDY APPLICANT: FIELD TITLE OF INVENTIO TITLE OF INVENTIO TITLE OF INVENTIO TITLE OF INVENTIO ADDRESSEE: JON STREE; GA COUNTRY: GAS ZIATE: GA TELETATION NUM PETELETATION NUM REGISTRATION NUM REGISTRATION TOR TELECATION TOWN TOWN TOWN TOWN TOWN TOWN TOWN TO	TEN SAN SAN SAN SAN SAN
90000000000000000000000000000000000000	RESULT US-08-9 Seque Paten GENE AP TITI TITI TITI TITI TITI TITI TITI T	FRAGE CORIG COR COR COR COLETY MA COLETY MA COLETY MA COLETY MA COLETY MA

us-09-491-146a-33.rai

```
NGS-08-635-886C-191
Sequence 191, Application US/08635886C
Patent No. 655114
GENERAL INFORMATION:
APPLICANT: LEROUX ROBLS, Geert
APPLICANT: DELEYS, Robert
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
FILE REFERENCE: 2752-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ouery Match 90.2%; Score 138; DB 4; Length 450; Best Local Similarity 89.3%; Pred. No. 1.2e-11; Matches 25; Conservative 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 90.2%; Score 138; DB 3; Length 28; Best Local Similarity 89.3%; Pred. No. 7.2e-13; Matches 25; Conservative 1; Mismatches 2: Indele
                                                                                                                                                         MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENY APPLICATION DATR:
APPLICATION NUMBER: US/08/921,887
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQKKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CURRENT APPLICATION NUMBER: US/08/635,886C CURRENT FILING DATE: 1996-04-25 PRIOR APPLICATION NUMBER: PCT/EP94/03555 PRIOR FILING DATE: 1994-00-28 PRIOR APPLICATION NUMBER: EP 93402718.6 PRIOR FILING DATE: 1993-11-04
                                                                                                                                                                                                                                                                                                                   CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REPERENCE/POCKET UMBER: 0306
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: not relevant MOLECULE TYPE: protein HYPOTHEICAL: NO ANTI-SPACE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PatentIn version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: hepatitis C virus
US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: not relev
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 internal
                                                                                     COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQ ID NOS: 286
SOFTWARE: Patentin version
SEQ ID NO 191
                                               Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FRAGMENT TYPE: 1
                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                             GENERAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOGALC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 25, Application US/08921887
Patent No. 603071
GENERAL INFORMATION:
APPLICANT: KHUDYAKO, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CARRENT APPLICATION DATA:
RILING DATE: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 96.7%; Score 148; DB 3; Best Local Similarity 96.4%; Pred. No. 2.9e-14; Matches 27; Conservative 0; Mismatches 1;
                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN WILLIAM L.
REGISTRATION NUMBER: 35,714
REPERENCE/DOCKET NUMBER: 03063-0380
TELECOMMUNICATION:
TELEPHONE: 404-818-3700.
                                                                                                                                                                          Sequence 28, Application US/08921887 Patent No. 6030771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: not relevant TOPOLOGY: not relevant MOLECULE TYPE: protein HYPOTHETICAL: NO ANTI-cent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEFAX: 404 818 3799
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                           Atlanta
: GA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FRAGMENT TYPE: 1
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEPHONE:
TELEFAX: 4
                                                                                                                                                       US-08-921-887-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-921+887-25
                                                                                                                               RESULT 2
                  ò
                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
```

ö

Gaps

```
Score 137; DB 3; Length 34;
Pred. No. 1.2e-12;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUW TYPE: MEDIUW MEDIUW TOO MEDIUM TOO MED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: YAMAGUCHI, Kenjiro
APPLICANT: KASHIWAKUMA, TOMIKO
APPLICANT: CHIBA, Yukie
APPLICANT: CHIBA, Yukie
APPLICANT: HASEGAWA, Akira
TITLE OF INVENTION: CHIMERA ANTIGEN PEPTIDE
NUMBER OF SEQUENCES: 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FRAGMENT TYPE: N-terminal ORIGINAL SOURCE: GRANISM: Human Hepatitis C Virus US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 36, Application US/09020846
Patent No. 6322965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STREET: BODG & LARDNER STREET: 3000 & Street, N.W. STATE: D.C
REFERENCE/DOCKET NUMBER: WP
TELECOMMUNICATION INFORMATION:
TELEPHAN: (703)836-6400
TELEPAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEO ID NO: 36
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 89.5%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                 SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
                                                                                                                                                       INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                  single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                  TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                          ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JS-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 6, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
PAPELICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                        Sequence 192, Application US/08635886C

Fatent No. 6555114:
GENERAL INFORMATION:
APPLICANT: LEROUX-ROELS, Geert
APPLICANT: LEROUX-ROELS, Geert
TITLE OF INVENTION: VIROS
FILE REFRENCE: 2752-18
CURRENT APPLICATION NUMBER: US/08/635,886C
CURRENT APPLICATION NUMBER: US/08/635,886C
FILE REFRENCE: 2752-18
CURRENT APPLICATION NUMBER: PCT/EP94/03555
FRIOR APPLICATION NUMBER: EP94-25
PRIOR APPLICATION NUMBER: EP93402718.6
PRIOR PELING DATE: 1994-10-28
PRIOR PELING DATE: 1993-11-04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     90.2%; Score 138; DB 4; Length 450; 89.3%; Pred. No. 1.2e-11; Live 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ZIP: USAN ZIP: USAN ZIP: ZIP: USAN ZIP: ZIP: ZIP: USAN ZIP: Eloppy disk MEDIUM TYPE: Floppy disk COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/380,160
                                                             PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-WAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OLIFF & BERRIDGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: Berridge, William P. REGISTRATION NUMBER: 30,024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOFTWARE: Patentin version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; TYPE: PRT; ORGANISM: hepatitis C virus
US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 89.3*
These 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET: P.O. Box 19928
CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQ ID NOS: 286
                                                                                                                                                                                                                 US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQ ID NO 192
LENGTH: 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-380-160-6
                             ò
                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      å
```

Gaps

Gaps

```
Sequence 1, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
APPLICANT: DALBON, Pascal
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: FOR DETECTING THE LATTER
UNDRES OF SEQUENCES:
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
                                                                                                                                                                                                                                                      Sequence 1, Application US/09389756

Patent No. 6576240

GENERAL INFORMATION:
APPLICANT: JOLIVET, MICHEL
APPLICANT: DALBON, PASCAL
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENI
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENI
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENI
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILE REFERENCE: 103959
CURRENT FILING DATE: 1999-09-07
EARLIER APPLICATION NUMBER: PCT/FR98/00442
BARLIER APPLICATION NUMBER: PCT/FR98/00442
BARLIER APPLICATION NUMBER: PCT/FR98/00442
BARLIER APPLICATION NUMBER: PCT/FR98/00442
SOFTWARRE: PATENTIN VOF: 2.1
SOFTWARRE: PATENTIN VOF: 2.1
                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CRANISM: Hepatitis C virus
CRGANISM: Hepatitis C virus
CRGANISM: Hepatitis C virus
FUBLICATION INFORMATION:
AUTHORS: Ogate N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
FITLE: Of Hepatitis Virus
COURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 189
PAGES: 3392-3396
DATE: 1991
CREEVANT RESIDUES: 2 TO 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                    Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2; Indels
                           Score 137; DB 3; L
Pred. No. 1.6e-12;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ch 89.5%; Score 137; DB 4; Similarity 89.3%; Pred. No. 1.6e-12; 25; Conservative 1; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1; Mismatches
                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                       1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
CITY: Alexandria
                      Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ZIP: 22320
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                  RESULT 9
US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE: V
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 10
US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
                                                                                                            ò
                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                              Sequence 2, Application US/08380160

Patent No. 6235284

GENERAL INFORMATION:
APPLICANT: DALBOW, PASCAL
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: POR DETECTING THE LATTER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSE:
ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
CITY Alexandria
                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                    ö
                                                                                                              Query Match
89.5%; Score 137; DB 4; Length 43;
Best Local Similarity 89.3%; Pred. No. 1.5e-12;
Matches 25; Conservative 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 22320

ZIP: 22320

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/NS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                  1 PKPÓRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                           5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: Beridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Human Hepatitis C Virus STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FRAGMENT TYPE: Noterminal ORIGINAL SOURCE: ORGANICE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEFAX: (703)836-2787
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCATION: 1.44
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE
                                                      ; MOLECULE TYF
US-09-020-846-36
                                                                                                                                                                                                                                                                                                          RESULT 8
US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-380-160-2
                                                                                                                                                                                                    à
                                                                                                                                                                                                                                        a
```

ö

Gaps

```
Sequence 23, Application US/08083947

Patent No. 5639594

GENERAL INFORMATION:
Patent No. 7639594

GENERAL INFORMATION:
TITLE OF INVENTION: In Diagnosing and Detecting No. 5639594-A, No. 5639594-1

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:
ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
CITY: Hauppauge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 137; DB 1; Length 61;
Pred. No. 2.2e-12;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 137; DB 1; Length 61; Pred. No. 2.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US/08/083,947
FILING DATE: 19930628
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 946,054
FILING DATE: 15-SEP-1993
ATTORNEY/AGENT INFORMATION:
      APPLICATION NUMBER: US/07/946,054
                                                                                                                       NAME: WILSON, M. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/DOCKET NUMBER: 2000
TELECOMUNICATION INFORMATION:
TELEPHONE: 516-273-2828
TELEFAX: 516-273-1717
INFORMATION FOR SEQ ID NO: 9:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34045
REFERENCE/DOCKET NUMBER: 20
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    89.5%;
                         15-SEP-1992
N: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (516)273-2828
                                                                                             ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Floppy disk
                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                         FILING DATE: 15
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: US
ZIP: 11.788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 9, Application US/07946054
Patent No. 5582968
GENERAL INFORMATION:
APPLICANT: Ward, Chang Y1
APPLICANT: Hosein, Barbara H
TITLE OF INVENTION: No. 5582968el Branched Hybrid and Cluster
TITLE OF INVENTION: No. 5582968 B Hepatitis
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: United Biomedical Inc.
STREET: 25 Davids Dr.
CITY: Hauppauge
STATE: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      89.5%; Score 137; DB 3; Length 45;
89.3%; Pred. No. 1.6e-12;
tive 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CTHER INFORMATION: /note= "N-terminal sequence of the OTHER INFORMATION: protein of the nucleocapside of the nucle
                                                   SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION UNMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM FC compatible
OPERATING SYSTEM: FC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                 CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: BETINGE, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ANTI-SENDS: "VERMINAL FRACKENT TYPE: N-terminal ORIGINAL SOURCE: ORIGINAL SOURCE: ORGANISM: Human Hepatitis C Virus
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0uery Match
Best Local Similarity 89.39
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CENGTH: 45 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: Hu
STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         03-07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
```

```
ATTORNEY/AGENT INFORMATION:
NAME: M. Lisa Wilson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEX: 421792,
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRANDEDNESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT-US93-08638-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 26, Application US/08262037
Patent No. 5747239
GENERAL INFORMATION
TITLE OF INVENTION: SYNTHETIC PEPTIDES SPECIFIC FOR
TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
TITLE OF INVENTION: INFECTION AND PREVENTION THEREOF AS VACCINES
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
        Gaps
                                                                                                                                                                           Sequence 3, Application US/08530550
Patent No. 5736321
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Wang, Chang Y1
TITLE OF INVENTION: Peptides Effective for Diagnosis and
TITLE OF INVENTION: Detection of Hepatitis c Infection
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89.5%; Score 137; DB 1; Length 61; 89.3%; Pred. No. 2.2e-12;
        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE: Floppy disk
COMFUTER: IBM PC compatible
COMPATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,550
    1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Mismatches
                                        1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                             4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/POCKET NUMBER: 2000
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVE.
                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
CITY: Hauppauge
25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Juery Match
est Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM: MEDIUM TYPE: Floppy of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEFAX: (516)273-171
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; MOLECULE TYPE: peptide US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                             USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ZIP: 11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-262-037-26
                                                                                                                                                            US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
Matches
                                                                                                                                           RESULT 13
                                        ò
                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
COMPATION TO THE PLOND LINE
COMPATION TO THE PLOND LINE
COMPATION SYSTEM: PC-DOSA'S-DOS
SOCTAMENT APPLICATION DAYS:
COMPATION SYSTEM: PC-DOSA'S-DOS
SOCTAMENT APPLICATION NUMBER: US/08/262.037
CURRENT APPLICATION NUMBER: US/08/262.037
PULING DAYE: US-08-1991
APPLICATION NUMBER: US/05/275
APPLICATION NUMBER: PC/1/59/20618
FULLOR DINNENTION: PC/1/59/20618
FULLOR DINNENTION: PC/1/59/20618
APPLICATION: WAND NUMBER: PC/1/59/20618
FULLOR DINNENTION: PC/1/59/20618
```

```
us-09-491-146a-33.rai
```

Search completed: August 7, 2003, 11:23:53 Job time : 10.6364 secs

ö

0; Gaps

```
August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                    al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                           153
1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                              451899 seqs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                              OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                    BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                      US-09-491-146A-33
                                                                                                                                                                                                      Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                      Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Database :
                                                                                                                                                                                                                                                                                                                                                  Searched:
                                                                                                                         Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

1. /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
2. /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
3. /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
3. /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
3. /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
3. /cgn2\_6/ptodata/1/pubpaa/PCUG\_PUBCOMB.pep:\*
3. /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
3. /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
3. /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
4. /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
5. /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
6. /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
7. /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
7. /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
7. /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
7. /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*

# SUMMARIES

		Description	Sequence 1, Appli	Sequence 10, Appl	Sequence 1, Appli	Sequence 8, Appli	Sequence 77, Appl	Sequence 14, Appl	Sequence 78, Appl	Sequence 46, Appl	Sequence 60, Appl		Sequence 152, App	Sequence 42, Appl	Sequence 44, Appl		Sequence 44, Appl	
SOMERICS		ID	US-10-367-677-1	US-09-851-138-10	US-09-758-308-1	US-09-756-875-8	US-09-921-397-77	US-09-851-138-14	US-09-921-397-78	US-09-851-138-46	US-09-851-138-60	US-09-899-046-152	US-09-878-281-152	. US-09-899-046-42	US-09-899-046-44	US-09-878-281-42	US-09-878-281-44	
		DB	15	10	σ	10	10	10	10	10	10	11	11	11	11	11	11	
	Query	Length	44	74	91	97	103	108	113	137	138	166	166	169	169	169	169	
æ	Query	Match	89.5	89.5	89.5	89.5	89.5	89.5	89.5	89.5	89.5	89.5	89.5	89.5	89.5	89.5	89.5	
		Score	137	137	137	137	137	137	137	137	137	137	137	137	137	137	137	
	Result	. o	1	7	e	₹	S	9	7	8	6	10	11	12	13	14	15	

89.5%; Score 137; DB 15; Length 44;

Query Match

Sequence 2, Appli Sequence 3, Appli Sequence 16, Appli Sequence 11, Appli Sequence 12, Appli Sequence 50, Appli Sequence 51, Appli Sequence 51, Appli Sequence 51, Appli Sequence 51, Appli Sequence 51, Appli Sequence 50, Appli Sequence 11, Appli Sequence 21, Appli Sequence 1, Appli Sequence 1, Appli Sequence 1, Appli Sequence 1, Appli Sequence 20, Appli	TIGENIC AND IMMUNOGENIC ING, PREVENTING AND TOT/FR98/00442 TOT/FR98/00442
US-09-929-955-2 US-09-194-966-2 US-09-181-138-19 US-09-851-138-19 US-09-851-138-18 US-09-851-138-18 US-09-899-046-52 US-09-899-046-54 US-09-899-046-144 US-09-899-046-144 US-09-878-281-54 US-09-973-025-50 US-09-973-025-50 US-09-973-025-50 US-09-973-025-50 US-09-973-025-50 US-09-973-025-50 US-09-971-611-23 US-09-975-80 US-09-975-97 US-09-975-97 US-09-975-97 US-09-975-95-11 US-09-975-97 US-09-975-97 US-09-975-97 US-09-975-97 US-09-975-97 US-09-975-97 US-09-975-97 US-09-975-97 US-09-975-97 US-09-975-97 US-09-975-97 US-09-975-97 US-09-97-97-97 US-09-97-97-97 US-09-97-97-97	ALIGNMENTS  A1  A1  C STRUCTURAL PEPTIDE, ANTIGENIC A  C STRUCTURAL PEPTIDE, ANTIGENIC A  DS, AND USES FOR DETECTING, PREVE  G AN HCV INFECTION  US/10/387,677  2-19  7.09/389,756  7.11ING DATE: 1998-03-05  ILING DATE: 1998-03-05  SC1. U.S.A.
182 10 3191 110 3191 110 3191 110 3191 110 3192 114 3193 119 3193 111 3193	ation US/1036 2030118604A1 2030118604A1 2030118604A1 2030118604A1 2030118604A1 20301186 20301018 20301018 20301018 20301018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 2031018 203108 203108 203108 203108 203108 203108 203108 203108 203108 203108 203108 203108 203108 203108 203108 203108 203108 203
œ œ œ œ œ œ œ œ œ œ œ œ œ œ œ œ œ œ œ	pplics 5. Colored Diversity of the property of
11111111111111111111111111111111111111	LLT 1  0-367-677-1  squence 1, Agenence 1,
	RESULT 1 US-10-367- Sequence; Publicat APPLICA
·	

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 97;
                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 8, Application US/09756875
Patent No. US20020150990A1
GENERAL INFORMATION:
APPLICANT: PIKE, IAN
TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 701-E Columbia Square
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 137; DB 10;
Pred. No. 2.7e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                   1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    E: Suite 701-E Columbia Square 555 13th Street, N. W.
                                                                                                                                                                                                                                                                                      1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: PCT/GB93/00410 FILING DATE: 26-FEB-1993 ATTORNEY/AGENT INFORMATION:
                           60/092,339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-AUG-1994
   CURRENT FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: 60/092,
PRIOR FILING DATE: 1999-07-10
NUMBER OF SEQ ID NOS: 5
SOCTWARE: Patentin version 3.0
SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
REFERENCE/FOCKEY NUMBER: 18
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
INFORMATION FOR SEQ ID NO: 8:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       89.5%;
                                                                                                                                                                                                               89.5%;
89.3%;
                                                                                                                                 TYPE: PRT
ORGANISM: Hepatitis C Virus
                                                                                                                                                                                                     Query Match
Best Local Similarity 89.38
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: unknown
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STREET: 555 13th sCITY: Washington STATE: D. C. COUNTRY: U. S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best_Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-756-875-8
                                                                                                                                                                          US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 1, Application US/09758308
Patent No. US20020090607A1
GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
FILE REFERENCE: 14114.0349U2
                                                                                                                                                                                                                                                         STUTVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
RILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: 08/836,075
FILING DATE: <URKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 74;
                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
 Pred. No. 1.2e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             89.5%; Score 137; DB 10; 89.3%; Pred. No. 2.1e-11;
                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Mismatches
                                                      1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                         TOPOLOGY: linear
MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CURRENT APPLICATION NUMBER: US/09/758,308
                                                                                                                                                            US-09-851-138-10; Application US/09851138; Sequence 10, Application US/09851138; Publication No. US20020183508A1; GENERAL INFORMATION: APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 74 amino acids
                                                                                                                                                                                                                                                                                                                     AGENTS
                                                                                                                                                                                                                                                                                                                                                                                           STREET: P.O. BOX 4433
CITY: HOUSTON
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             14114.0349U2
                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 89.39
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                STATE: TEXAS
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 3
US-09-758-308-1
                                                                                                                                               RESULT 2
                                                                                     셤
                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

ô

```
APPLICANT: MAERTENS, GEERT
STUYVER, LIEVENEN
TITLE OF INVENTION - NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGM
                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                    Sequence 78, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and TITLE OF INVENTION: applications thereof File Reference: adaptations thereof File Reference: AJAZ
CURRENT FILING DATE: 2009/991397
PRIOR FILING DATE: 2000-08-03
PRIOR FILING DATE: 2000-08-03
                                                                                                             ·.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels 0;
                                                                      Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: USA
ZIP: 77210-4433
COMPUTER READBALE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 137; DB 10;
Pred. No. 3.2e-11;
1; Mismatches 2;
                                                                      DB 10;
                                                                  Ouery Match 89.5%; Score 137; DB 10; Best Local Similarity 89.3%; Pred. No. 3.1e-11; Matches 25; Conservative 1; Mismatches 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
                                                                                                                                                         1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                1 PKPORKTKRKAHRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; SEQUENCE DESCRIPTION: SEQ ID NO: 14: US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,075
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 46, Application US/09851138 Publication No. US20020183508A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQ ID NOS: 156
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-851-138-46
                                                                                                                                                                                                                                                                                         US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH: 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                         á
                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HEPATITIS C VIRUS GENOTYPES
PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                  Sequence 77, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
APPLICANT: HYBRIGGEN
TITLE OF INVENTION: Dathogenic acids and polypeptides selected from TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and TITLE OF INVENTION: applications thereof FILE REFERENCE: B4809A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2000-08-02
PRIOR FILING DATE: 2000-08-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY: USA
ZIP: 77210-44
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
COMPUTER: MISTEM: PC-TOS/MS-DOS
OFFWARE: MISTOSOFT Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 137; DB 10;
Pred. No. 2.9e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: 08/836,075
FILING DATE: CUNKNOWN>
APPLICATION NUMBER: EP 94870166.9
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. 2.9e
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REGISTRATION NUMBER: 29,775
REFERRENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF
AND THEIR USE AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 14, Application US/0951138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    89.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: Hepatitis C virus --09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                      NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 77
LENGTH: 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 6
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Juery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

Fri Aug

```
ö
                                                                                                                                                                                                                                                                              New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT:
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy, NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                     TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylax. NUMBER OF SEQUENCES: 270

WOMPUTER READBLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SOFTWARE: Tatentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89.5%; Score 137; DB 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No. 4.76
1; Mismatches
                                                                             1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 137;
                                                                                                                                                                         US-09-899-046-152; Sequence 152. Application US/09899046; Publication No. US20030008274A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 152, Application US/09878281.
Publication No. US20030032005A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : 166 amino acids
amino acid
              25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
              Matches
                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                Length 137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURREWT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                          Score 137; DB 10;
Pred. No. 3.9e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 137; DB 10;
Pred. No. 3.9e-11;
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                        ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433 CITY: HOUSTON STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                         TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; Sequence 60, Application US/09851138; Publication No. US20020183508A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: FLORPY disk
                                                                                                                                                        SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                          89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES: 20'
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 89.37
"---hes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
                                                                                                                                                                                                                                                                                  US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                               3
```

Length 166;

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                      Gaps
                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 169;
         Length 169
                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/878,281
  Score 137; DB 11;
Pred. No. 4.8e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 137; DB 11;
Pred. No. 4.8e-11;
1; Mismatches 2;
                                                                                        1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                   5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                       Sequence 42, Application US/09878281 Publication No. US20030032005A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 44, Application US/09878281
Publication No. US20030032005A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89.5%;
ilarity 89.3%;
Conservative 1
  Query Match 89.5%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                        Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : 169 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           169 amino acids amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               , MOLECULE TYPE: protein US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; MOLECULE TYPE:
US-09-878-281-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                RESULT 14
US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-878-281-44
                                                                                          ô
                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
COMPUTER: IBM PC COMPATION POSSAMS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                      TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
                           Gaps
                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 11; Length 169;
  Pred. No. 4.7e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERTING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89.5%; Score 137; DE 1larity 89.3%; Pred. No. 4.8e Conservative 1; Mismatches
                                                               1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                        5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                Sequence 42, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 44, Application US/09899046 Ublication No. US20030008274A1 GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
INFORMATION FOR SEQ ID NO: 42:
88.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
INFORMATION FOR SED ID NO: 4
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acids
                      25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 25; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
  Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY:
                                                                                                                                                                                          US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-899-046-44
                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
```

```
Query Match
Best Local Similarity 89.3%; Score 137; DB 11; Length 169;
Best Local Similarity 89.3%; Pred. No. 4.8e-11;
Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

δ Q

Search completed: August 7, 2003, 12:01:15 Job time: 15.3636 secs

S FARUKKIKKNINKRPODVKFPGGGOIVG 32

Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

Searched:

Database

Result

Š

450/00

```
Hepatitis C virus
NC mosaic protein
Hepatitis C virus
Blood transmiscibl
HCV-SI full-length
Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
Non-A, Non-B Hepat
Non-A, Non-B Hepat
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Non-A, Non-B Hepa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NC mosaic protein amino acid fragment J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                                                                                                                                                                                                                                                                                                                                                                  AAW66083
AAY26952
                                                                                                                                                                                                                                                                                                                                                       AAR54065
AAW37380
                                                                                                                                                                                                                                                                                                                                                                                                                 AAY94410
AAY94409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY06682 standard; Protein; 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98WO-US17385.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            17-JUN-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Fields HA, Khudyakov YE;
                                                                                      470
28
3010
3010
                                                                                                                                                                            WPI; 1999-204671/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
 25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              409910506-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          04-MAR-1999.
AAY06682;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY06682
   NC mosaic protein
Hepatitis C virus
Encoded by Hepatit
Encoded by Hepatit
Encoded by Hepatit
NC mosaic protein
HCV type 3 capsid
Hepatitis C virus
Hepatitis C virus
                                                                                                 Search time 38.5455 Seconds (without alignments)
115.301 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                   1. SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*

SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*

SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*

SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*

SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*

SIDSI/gcgdata/geneseqy-embl/AA1982.DAT:*

SIDSI/gcgdata/geneseqy-embl/AA1986.DAT:*

SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:*

SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:*

SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*

SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /SIDS1/gcgdata/geneseq/geneseqp·embl/AA2001.DAT:*/SIDS1/gcgdata/geneseq/geneseqp·embl/AA2002.DAT:*/SIDS1/gcgdata/geneseq/geneseqp·embl/AA2003.DAT:*
            GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                        .al number of hits satisfying chosen parameters:
                                                                                                                                                                       149
1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                             1107863 segs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                 August 7, 2003, 11:05:37;
                                                                                                                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                       protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR34473
AAR34474
AAR34475
AAY06673
AAB71258
AAR92968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY06682
AAR94462
                                                                                                                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                          US-09-491-146A-32
                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20
114
120
171
171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3023
470
470
470
189
191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
Match Length
```

셤

ò

```
Gaps
                                                                                                                                                                     Gaps
(HCV) genomic RNA. The protein can be easily detected by antibodies in an assay for the detection of HCV. The DNA and the protein are useful for classifying the subtype of HCV. At least a part of the protein may be used as a vaccine against HCV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cDNA was prepared from HCV genomic RNA. Full-length clone JKI-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-A. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                 ô
                                                                                                                                Length 3023;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 470;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    non-B hepatitis virus; NANBHV; liver disease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNA and cDNA of hepatitis C virus – useful as probes for
diagnosing HCV infection
                                                                                                                            100.0%; Score 149; DB 17;
100.0%; Pred. No. 7e-12;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    87.2%; Score 130; DB 14;
82.1%; Pred. No. 4.8e-10;
live 4; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polymerase chain reaction; diagnostic method
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Encoded by Hepatitis C virus clone JK3-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                      1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                         5 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 4; Page 26-28; 44pp; Japanese.
                                                                                                                                                                                                                                                                                                                                      Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (SANW ) SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                                  AAR34473 standard; Protein; 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR34474 standard; Protein; 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91JP-0153736.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91JP-0153736.
                                                                                                       Query Match

Query Match

Best Local Similarity

Local Similarity

Local Similarity

Local Similarity

Local Similarity

Local Similarity

Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                          30-JUL-1993 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30-JUL-1993 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 82.1'
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                See AAQ40425-Q40439.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1993-130638/16.
                                                                                         3023 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAQ40431.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JP05068562-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23-MAR-1993.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR34474;
                                                                                                                                                                                                                                                                                                                                                                       AAR34473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Seguence
                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 4
AAR34474
ID AAR3
XX
AC AAR3
XX
DT 30-J
                                                                                                                                                                                                                                                                                               RESULT 3
                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                      a
8 X C C C C C
                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                              The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (RRAL). The mosalc protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme Sequences Asv06673-681 represent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence is a polypeptide comprising a 3023 amino acid sequence which is encoded by a cDNA (AAT13279) to a hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus genomic RNA, DNA and related proteins - useful for detection, diagnosis and identification of hepatitis C virus sub-type
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1505..1520
/note= "this part of the sequence is missing from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "this part of the sequence is missing from the specification"
                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                     Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hepatitis C virus; antibody; detection; diagnosis; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                   Score 149; DB 20;
Pred. No. 5.2e-14;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              the specification"..2448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 4; Page 16-23; 25pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR94462 standard; Protein; 3023 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                   100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus polypeptide.
               Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94JP-0223933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94JP-0223933
                                                                                                                                                                                                                                                                                                                                                          very Match
st Local Similarity 100.0
witches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1996-182301/19.
N-PSDB; AAT13279.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 subtype.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (SAYA/) SAYAMA K.
                                                                                                                                                                                                                                                                                                                                  28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JP08056672-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20-SEP-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               05-MAR-1996.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 classify;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR94462;
                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR94462
ID AAR9
```

ö

Gaps

ö

Indels

28

Length 470;

```
Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                              CDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-C. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV. See AAQ40425-Q40439
                                                                                                                    Score 130; DB 14;
Pred. No. 4.8e-10;
4; Mismatches 1;
                                                                                                                                                                            (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   84.6%; Score 126;
                                                                                                                                                                1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG
                                                                                                                                                                                                                                                                                                          mosaic protein amino acid fragment
          Claim 4; Page 30-32; 44pp; Japanese.
                                                                                                                                                                                                                                           Ā
                                                                                                                                                                                                                                          AAY06673 standard; Protein; 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 5; Fig 9; 66pp; English
                                                                                                                    Query Match
Best Local Similarity 82.1%;
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                               98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                     97US-0921887
                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Khudyakov
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1999-204671/17;
                                                                                                 470 AA;
                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28 AA;
                                                                                                                                                                                                                                                                                                                                                                                   WO9910506-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                               21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-AUG-1997;
                                                                                                                                                                                                                                                                                      17-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                          04-MAR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Fields HA,
                                                                                                 Sequence
                                                                                                                                                                                                                                                               AAY06673;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                * 8 × 5 5 5 5 5 5 5 × 8
                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                    cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-B. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV.
                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                          Length 470;
                              NANBHV; liver disease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     non-B hepatitis virus; NANBHV; liver disease;
                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA and cDNA of hepatitis C virus – useful as probes for diagnosing HCV infection
                                                                                                                                                                                                                         DNA and cDNA of hepatitis C virus - useful as probes for diagnosing \ensuremath{\mathsf{HCV}} infection
                                                                                                                                                                                                                                                                                                                                                                         Score 130; DB 14;
Pred. No. 4.8e-10;
4; Mismatches 1;
                              HCV; non-A, non-B hepatitis virus; NANBHV; l
polymerase chain reaction; diagnostic method
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polymerase chain reaction; diaqnostic method
                                                                                                                                                                                                                                                                                                                                                                                                                   Encoded by Hepatitis C virus clone JK3-B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Encoded by Hepatitis C virus clone JK3-C.
                                                                                                                                                                                                                                                                Claim 4; Page 28-30; 44pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ź
                                                                                                                                                                          (SANW ) SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (SANW ) SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR34475 standard; Protein; 470
                                                                                                                                                                                                                                                                                                                                                                          87.2%;
82.1%;
                                                                                                                                 91JP-0153736.
                                                                                                                                                    91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91JP-0153736
                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 82.1",
Associated 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1993-130638/16
                                                                                                                                                                                                WPI; 1993-130638/16
                                                                                                                                                                                                                                                                                                                                See AAQ40425-Q40439
                                                                                                                                                                                                                                                                                                                                                    470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAQ40433,
                                                                                                                                                                                                           N-PSDB; AAQ40432
                                                                                    JP05068562-A.
                                                                                                                                30-MAY-1991;
                                                                                                                                                   30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30-JUL-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JP05068562-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30-MAY-1991;
                                                                                                          23-MAR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV; non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23-MAR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR34475;
                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR34475
ID AAR3
XX
74×8×800000×8
                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
```

```
The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from uncleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein a laso useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, specitum of immunoreactivity, and antigen specificity of enzyme sequences AAV66673-681 represent amino acid sequence of each monomer comprising the NC mosaic protein.
           antigenic
for
New mosalc protein, comprising a plurality of homologous peptides from different genotypes of a species - useful detecting hepatitis infection in an individual
```

Length 28;

DB 20;

셤

ò

```
HCV; El; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV; El; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated CDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the CDNAs can be used in vaccines for immunishing against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, Lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection \,
                                                                                                                                                                                                             Hepatitis C virus isolate HK10 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus isolate S52 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG
                                                                                     AAR92968 standard; Protein; 191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR92969 standard; Protein; 191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 4; Page 207; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Purcell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 78.6%;
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                         95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94US-0290665
                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              02-OCT-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1996-139709/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              191 AA;
                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; AAT16642
                                                                                                                                                                                                                                                                                                                                                        409605315-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .5-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                         .5-AUG-1995;
                                                                                                                                                                      02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                22-FEB-1996.
                                                                                                                                                                                                                                                                           hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       hepatitis.
                                                                                                                           AAR92968;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR92969;
                                            RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR92969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This invention describes a novel attenuated flavivirus live vaccine comprising a flavivirus mutant that has a deletion of at least 4 consecutive anno acids from the capsid protein, provided that the consecutive anno acids from the capsid protein, provided that the consecutive annoal hydrophobic region is not affected by the deletion. The vaccine of the invention has virucide, antinflammatory and hepatotropic activity. The attenuated vaccine, and similar nucleic acid vaccines that encode the mutated capsid protein, are useful for protection against a wide range of flavivirus diseases, e.g. yellow fever. Japanese and concephalitis, dengue, classical swine fever, bovine viral diarrhoea and concephalitis C. The specified deletion: (i) produces a reliably attenuated virus that does not revert to virulence; (ii) is exactly defined and does: (c) the step in much responses to important proteins; and (ii) can not generate a non-natural virus by recombination. The mutant viruses eliminate the need to produce large amounts of infectious/virulent contracts and can be produced with less expense. The protective response continues, and can be produced with less expense. The protective response continues are presents a fragment of the capsid protein from Hepatitis
                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Attenuated flavivirus live vaccine, useful for protection against e.g. yellow fever, comprises virus with attenuating deletion of amino acids from the capsid protein
                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     c virus (HCV) type 1 described in the disclosure of the invention
                                                                                                                                                                                                                                                                                                                                                                     Capsid protein; attenuated vaccine; virucide; antinflammatory; hepatotropic; yellow fever; Japanese encephalitis; dengue; classical swine fever; bovine viral diarrhoea; hepatitis C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 189;
                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
Best Local Similarity 78.6%; Pred. No. 9.1e-11;
Matches 22; Conservative 5; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11arity 78.6%; Score 126; DB 23; 78.6%; Pred. No. 6.8e-10; Conservative 5; Mismatches 1;
                                                         1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                               PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Š
                                                                                                                                                                                                                                                                                                                                HCV type 3 capsid protein fragment.
                                                                                                                                                                                                      AAB71258 standard; protein; 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Fig 2; 30pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11-FEB-2002; 2002WO-AT00046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21-FEB-2001; 2001AT-0000272.
                                                                                                                                                                                                                                                                                   18-NOV-2002 (first entry
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2002-667064/71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Heinz FX, Mandl C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 22; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (HEIN/) HEINZ F X. (MAND/) MANDL C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      depatitis c virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            189 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200266621-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29-AUG-2002.
                                                                                                                                                                                                                                               AAB71258;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Seguence
                                                                                                                                                                                 RESULT 7
```

388888888888888

ò 셤

RH;

ö

Gaps

ö

Length 191; Indels

Score 126; DB 17; Pred. No. 6.8e-10; 5; Mismatches 1;

S

ö

us-09-491-146a-32.rag

3888888888

```
o acid in this position is designated X in specification, but codon usage shows that the possible amino acid at this pos. is Gly"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid in this position is designated X in the specification, but codon usage shows that the only possible amino acid at this pos. is Gly"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus; subtype; polymerase chain reaction; amplification; PCR; primer; probe; antibody; infection.
                                                              AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other monounclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                Gaps
      to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus types 7c(8a) isolates VN4 amino acids 1-317.
                                                                                                                                                                                                                                               Length 191;
                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                         Score 126; DB 17;
Pred. No. 6.8e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /label- Gly
/note- "amino acid in this po
the specification, b
only possible amino
                                                                                                                                                                                                                                                            Pred. No. 6.86
5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note= "amino acid in this
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              _{\rm G1y}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Met, Thr, Lys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    'label- Val, Ala, Glu, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ser, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ile, Val
                                                                                                                                                                                                                                                                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                           Misc-difference 144..149
/label= Val, Ala, Glu,
                                  Claim 4; Page 209-210; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       label- Val, Ala, Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                     AAR96547 standard; peptide; 319 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               label Cys, Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Leu
                                                                                                                                                                                                                                          94.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              'label - Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'label- Ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /label= Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /label- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /label= Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                         Local Similarity 78.6 les 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 233
                                                                                                                                                                                                             191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10-MAR-1997
                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR96547;
                                                                                                                                                                                                                                          Query Match
      nseq
                                                                                                                                                                                                                                                              Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                   AAR96547
      a
                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                         AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                 DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA and amino acid sequence of HCV envelope 1 and core proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 126; DB 17; Length 191;
Pred. No. 6.8e-10;
5; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus isolate DK12 core protein,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                               Claim 4; Page 208; 340pp; English
                                                                                                                                                                     Miller RH, Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR92971 standard; Protein; 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           84.6%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Miller RH, Purcell
                                                           95WO-US10398
                                                                                          94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             02-OCT-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                                   WPI; 1996-139709/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NPI; 1996-139709/14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                                                                   N-PSDB; AAT16643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAT16645
WO9605315-A2
                                                                                          15-AUG-1994;
                                                           15-AUG-1995;
                             22-FEB-1996.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22-FEB-1996,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR92971;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; El;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bukh J,
                                                                                                                                                                     Bukh J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
AAR92971
ID AAR92
XX
AC AAR92
```

g ô

09-MAY-1996

Maertens G,

```
The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from uncleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein a laso useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunozactivity, and antigen specificity of enzyme immunosassys. This provides improved detection of hepatitis C virus. Sequences AAV06673-681 sepresent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus; subtype; polymerase chain reaction; amplification; PCR; primer; probe; antibody; infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus poly:nucleic acid unique to unidentified sub:type - used to develop probes and primers for new sub:types and vaccines to prevent and treat infection
                                                                                                                                                    New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus type 11a isolate FR19 amino acids 1-74.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 125; DB 20;
Pred. No. 1.3e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    83.9%; Scor.
82.1%; Pred. No. 1..
3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                           (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR96577 standard; peptide; 74 AA.
                                                                                                                                                                                                                                          Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             95EP-0870076.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    95WO-EP04155.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 82.1
Matches 23; Conservative
                                                                   Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (INNO-) INNOGENETICS NV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Maertens G, Stuyver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1996-251460/25.
N-PSDB; AAT27988.
                                                                                                            WPI; 1999-204671/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                23-OCT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO9613590-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-JUN-1995;
21-OCT-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14-MAR-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09-MAY-1996.
                                                               Fields HA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR96577;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR96577
      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The sequences AAR96526-R96578 represent novel sequences isolated from hepatitis C virus subtypes different from subtypes la-c, 2a-d, 3a-f, 4a-j, 5a and 6a. They esp. from the novel subtypes ld-f, 2e-i, 2k, 2l, 3g, 4k-m, 7a-c or types 9, 10 or 11. The sequences corresp. to the 5 unitranslated region (UR), the Core/El, NS4 or NS5B regions of the genome. This sequence represents amino acids 1-317 from the HCV types 7c and 8a isolates VN4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The new HCV types were isolated from patients with chronic HCV from the benelux countries. France, Cameroon and Vietnam, because of their aberiant reactivities. The RNA was extracted, cDNA synthesised and PCR amplified, cloned and genotyped. The YUR, Core/El and NS5B regions were ageinence either directly or partially and used to classify the new viruses into (sub)types based on comparison with known sequences. The sequences were used to generate the peptides AAR96424-R96524. The aequences can also be used to synthesise probes and primers for the detection of HCV in a sample. The polypeptides can be used codetect anti-HCV antibodies, for HCV typing or to prevent HCV infections.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus poly:nucleic acid unique to unidentified sub:type - used to develop probes and primers for new sub:types and vaccines to prevent and treat infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         84.6%; Score 126; DB 17; Length 319; 78.6%; Pred. No. 1.2e-09; ive 5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 restriction endonuclease assisted ligation; vaccination
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NC mosaic protein amino acid fragment B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY06674 standard; Protein; 28 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 25; Fig 3; 150pp; English
                                                                                 95WO-EP04155
                                                                                                                         95EP-0870076
94EP-0870166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                          (INNO-) INNOGENETICS NV
                                                                                                                                                                                                                                     Stuyver L;
                                                                                                                                                                                                                                                                           WPI; 1996-251460/25.
N-PSDB; AAT27958.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   319 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
WO9613590-A2
                                                                                                                         28-JUN-1995;
21-OCT-1994;
                                                                                 23-OCT-1995;
```

Sequence

85888888888888888888888888888888

Ouery Match

õ g 17-JUN-1999

AAY06674;

hesULT 12 AAY06674

21-AUG-1998; 25-AUG-1997;

04-MAR-1999

against HCV in serum, The antibodies may

ö

Gaps

ö

Indels

Length 191;

```
HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, sallva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection \,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 125; DB 17; Length 191;
Pred. No. 9.5e-10;
6; Mismatches 1; Indels
                                                                                                                        Score 125; DB 17;
Pred. No. 9.5e-10;
3; Mismatches 2;
proteins may also be used to detect antibodies a saliva, lymphocytes or other mononuclear cells. be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus isolate S2 core protein.
                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                           1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US DEPT HEALTH & HUMAN SERVICES. US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Search completed: August 7, 2003, 11:14:10
Job time : 39.6364 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR92970 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 4; Page 209; 340pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      83.9%;
75.0%;
                                                                                                                     Query Match
Best Local Similarity 82.1%;
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 83.9
Best Local Similarity 75.0
Matches 21; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1996-139709/14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             191 AA;
                                                                                  Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAT16644
                                                                                191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9605315-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                          02-0CT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22-FEB-1996.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                    AAR92970;
                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (USSH)
                                                                                                                                                                                                                                                                                                    RESULT 15
AAR92970
                                                                                                                                                                                                                                                                                                                                               ខ្លួនទ
                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                           The new HCV types were isolated from patients with chronic HCV from the Benelux countries, France, Cameroon and Vietnam, because of their Benelux countries, France, Cameroon and Vietnam, because of their aberrant reactivities. The RNA was extracted, cDNA synthesised and PCR amplified, cloned and genotyped. The 5'UR, Core/El and NS5B regions were sequenced atther directly or partially and used to classify the new viruses into (sub)types based on comparison with known sequences. The sequences can also be used to generate the peptides ARP96424-R96524. The sequences can also be used to synthesise probes and primers for the detection of HCV in a sample. The polypeptides can be used to detect anti-HCV antibodies, for HCV typing or to prevent HCV infections.
                                              The sequences AAR96526-R96578 represent novel sequences isolated from hepatitis C virus subtypes different from subtypes la-c, 2a-d, 3a-f, 4a-j, 5a and 6a. They esp. from the novel subtypes ld-f, 2e-i, 2k, 2l, 3g, 4k-m, 7a-c or types 9, 10 or 11. The sequences corresp. to the 5' untranslated region (UR), the Core/El, NS4 or NS5B regions of the genome. This sequence represents amino acids 1-74 from the HCV type lla isolate FR19.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. Of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     envelope 1; core protein; HCV genotyping; antibody; vaccine,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 125; DB 17;
Pred. No. 3.5e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus isolate S45 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (USSH ) US DEPT HEALTH & HUMAN SERVICES (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 4; Page 186-187; 340pp; English.
                Claim 25; Fig 3; 150pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR92943 standard; Protein; 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                            83.9%;
82.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       02-OCT-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1996-139709/14.
N-PSDB; AAT16617.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                        74 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22-FEB-1996.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR92943;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 14
```

ž

à ద

ö

Gaps

ö

8 09:20:26 2003

Fri Aug

```
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

US-09-491-146A-32 149 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283308 seqs, 96168682 residues Searched:

.al number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

PIR\_76:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

	Description	genome polyprotein								genome polyprotein								genome polyprotein								genome polyprotein			genome polyprotein	genome polyprotein
SUMMARIES	<b>a</b>	S19875	PC2060	S41359	S41358	S41360	PC2061	S41288	S41351	S41349	JQ1584	S41356	A45573	S21336	S41361	S41353	S41355	S41357	541348	S41371	S41341	S41370	S41369	S41368 ·	S41342	34	35	ß	m	S41347
	DB	7	7	~	N	ď	~	7	~	7	N	~	~	~	7	N	~	~	7	~	7	~	N	7	N	N	~	~	~	7
	Length	782	411	114	114	124	411	492	115	115	640	108	3010	88	123	108	108	108	108	112	112	114	114	114	115	115	115	115	115	115
4	Ouery Match				84.6										•				•	•	•	•			•	•		•		•
	Score	130	129	126	126	126	126	126	125	125	125	123	123	122	122	121	121	121	121	121	121	121	121	121	121	121	121	121	121	121
	Result No.	н	7	m	4	'n	9	7	œ	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	53	24	25	56	27	28	29

RESULT 2 PC2060

Protone polyprotein N1 - hepatitis C virus
N; Contains: envelope protein E1; nonstructural protein E2/NS1
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Species: pepatitis C virus
S; Vitvitski, L.; Tong, S.P.; Trepo, C.
Blochem: Biophys. Res. Commun. 199, 1474-1481, 1994
A; Title: Identification of the third major genotype of hepatitis C virus in France
A; Reference number: PC2060; MUID: 94197744; PMID: 8147893
A; Accession: PC2060
A; Molecule type: mRNA
A; Residues: 1-411 - 6L1J>
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP; capabid protein; envelope protein; glycoprotein; nonstructural protein
E; 132-383/Product: envelope protein E1 #status predicted <NPE>
F; 196, 209, 234, 305, 325/Binding site: carbohydrate (Asn) (covalent) #status predicte

genome polyprotein genome polyprotein genome polyprotein genome polyprotein genome polyprotein polyprotein polyprotein polyprotein - hepa genome polyprotein genome polyprotein genome polyprotein hypothetical prote genome polyprotein	ALIGNMENTS	SULT 1  Contains: core protein - hepatitis C virus (isolate JK3) (fragment)  Contains: core protein; envelope protein 1; nonstructural protein 2; NSI/E2 prot Species: hepatitis C virus Variety: isolate JK3 Variety: Murakami, S. Honda, M.; Kaneko, S.; Massshi, U.; Kobayashi, K.; Murakami, S. Honda, M.; Kaneko, S.; Massshi, U.; Kobayashi, K.; Murakami, S. Honda, M.; Kaneko, S.; Massshi, U.; Kobayashi, K.; Murakami, S. Honda, M.; Kaneko, S.; Massahi, U.; Kobayashi, K.; Murakami, S. Honda, M.; Kaneko, S.; Massahi, U.; Kobayashi, K.; Murakami, S. Honda, M.; Kaneko, S.; Massahi, U.; Kobayashi, K.; Murakami, S. Honda, M.; Kaneko, S.; Massahi, U.; Kobayashi, K.; Murakami, S. Molecule type: genomic RNA Residues: 1.782 - KNO Residues: 1.782 - KNO Residues: I.782 - KNO Residues: I.783 - Murakami, S. I.91.783 - Murch I.91.782 - Murch I.91.782 - Murch I.91.782 - Murch I.91.782 - Murakanis Residues I.91.783 - Murch I.91.783 - Murch I.91.782 - Murakanis Residues I.91.783 - Murch I.91.784 - Murch I.91.785 - Murch I.91.785 - Murch I.91.786 - Murch I.91.786 - Murch I.91.787	
\$41343 \$214346 \$214346 \$12707 \$01026 \$01026 \$19071 \$18032 \$00883 \$00883 \$00883	ALIG	roctein - hepatitis C virus (isolicore protein; and envelope protein; lepatitis C virus lisolate JK3 lisolate JK3 lisolate JK3 lisolate JK3 lisolate JK3 Kaneko, S.; Masashi, U.; Kobayas Ni. Sequence analysis of putative number: S18029 S19875 S1987 S19875	
напапапапапапапан		intis C envelop insashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, labrashi, l	
1115 369 369 523 523 782 782 782 782 782 782 782 783		ein - hepatitis C virus titis C virus ate JK3 ate EMBL bar Library, Sep Sequence analysis of put bber: S18029 sequence analysis of put sicore protein status per cic: nonstructural protein core protein status fc: nonstructural protein conservative 4; Mi kPKRROTKRNTLRRPKNVKFPAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	
		oothouse of the control of the contr	
		rotein -   core protein	
		RESULT 1 S19875  Senome polyprotein - hepatitis C vii N;Contains: core protein; envelope i N;Contains: core protein; envelope i N;Contains: core protein; envelope i N;Catety: isolate JK3 C;Date: 30-010-1992 #sequence_revis: C;Accession: S19875 R;Honda, M.; Kaneko, S.; Masashi, U submitted to the EMBL Data Library, A;Description: Sequence analysis of A;Description: S19875 A;Accession: S19875 A;Accession: S19875 A;Accession: S19875 A;Accession: S19875 A;Accession: S19875 A;Accession: S1987 A;Coss-references: EMBL:X61592; NII A;Coss-references: EMBL:X61592; NII A;Coss-references: EMBL:X61592; NII A;Coss-references: capaid protein; core protein F;191-Product: core protein #stati F;191-Product: core protein #stati F;191-Product: core protein #stati F;191-Product: nonstructural protein F;344-782/Product: nonstructural productes Agency Match Best Local Similarity 82.1%; St Best Local Similarity 82.1%; St HII:	
00000000000444444000000000000000000000		RESULT 1 S19875 genome polyprotein - hepatitis C virus ( genome polyprotein - hepatitis C virus ( N;Contains: core protein; envelope protein; species: hepatitis C virus A;Varlety: isolate JK3 B;Honda, M.; Kaneko, S.; Masashi, U.; Ko B;Description: Sequence analysis of puta A;Reference number: S18029 A;Reference number: S18029 A;Reference number: S18029 A;Residues: 1-782 <-HON> A;Residues: 1-782 <-HON> A;Residues: 1-782 <-HON> A;Residues: 1-782 <-HON> A;Reperimental source: isolate JK3 C;Superfamily: hepatitis C virus genome F;1-191/Product: core protein #status F;134-782/Product: core protein #status F;34-733/Product: nonstructural protein F;134-782/Product: nonstructural protein F;134-782/Product: nonstructural protein F;134-782/Product: nonstructural protein F;134-782/Product: nonstructural protein F;134-181/Fill	

```
541288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Guoume purprotects. ...epatilis viius (genotype 5, N1) (Iragiment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety; genotype 3, N1
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: 541358
...van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
...iited to the EMBL Data Library, January 1994
Ascription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41318
A;Accession: 841358
A;Accession: 841358
A;Accession: 841358
A;Accession: 841358
A;Accession: 84158
A;Accession: 84158
A;Cross-references: EMBL:229461; NID:g443884; PIDN:CAA82599.1; PID:g443885
A;Experimental source: genotype 3, N1
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <MAT>
                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                  N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotypes 3, N2
A.Variety: genotypes 3, N2
C.Sacession: S41359
R.Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
Secription: Analysis of hepatitis C virus genotypes 1 to 5 by Lipa.
                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Cross-references: EMBL:229462; NID:g443886; PIDN:CAA82600.1; PID:g443887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                             genome polyprotein - hepatitis C virus (genotype 3, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein - hepatitis C virus (genotype 3, N1) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    S41360
genome polyprotein - hepatitis C virus (genotype 3, N3) (fragment)
    Length 411;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 114;
                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A:Experimental source: genotype 3, N2
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; polyprotein
F:1-114/Product: core protein *status predicted <MAT>
Score 129; DB 2;
Pred. No. 3.1e-10;
5; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match

84.6%; Score 126; DB 2;
Best Local Similarity 78.6%; Pred. No. 2.3e-10;
Matches 22; Conservative 5; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     84.6%; Score 126; DB 2; 78.6%; Pred. No. 2.3e-10; ive 5; Mismatches 1
                                                                                                                1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          86.6%;
78.6%;
                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 84.6
Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: genomic RNA A: Residues: 1-114 <VAN>
                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Accession: $41359
                     Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                      ò
                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
```

```
genome polyprotein N2 - hepatitis C virus
N;Contains: envelope protein E1; nonstructural protein E2/NS1
N;Contains: envelope protein E1; nonstructural protein E2/NS1
C;Species: hepatitis C virus
C;Species: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Nov-2000
C;Accession: PC2061
R;Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
R;Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
A;Itle: Identification of the third major genotype of hepatitis C virus in France.
A;Reference number: PC2060; MUID:94197744; PMID:8147893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        prot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; envelope protein; nonstructural protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A.Cross-references: GB:L12355, NID:9410169; PIDN:AAA.20155.1; PID:9410170
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capald protein; envelope protein; glycoprotein; nonstructural
F;192-383/Product: envelope protein E1 #status predicted <SPE>
F;384-411/Product: nonstructural protein E2/NS1 #status predicted <NPE>
F;196,209,234,305,325/Binding site: carbohydrate (Asn) (covalent) #status predi
A.Variety: denotype 3, N3
C.Species: hepatitis C virus
A.Variety: genotype 3, N3
C.Sate: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C.Saccession: $41360
R.Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A.Reference number: $41341
A.Reference num
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (fragment)
N;Contains: core protein; envelope protein; NS1 protein
C;Species: hepatitis C virus
C;Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A)Cross-references: EMEL:229463; NID:9443888; PIDN:CAA82601.1; PID:9443889
A)Experimental source: genotype 3, N3
C;Superfamily: hepartitis C virus genome polyprotein
C;Keywords: capaid protein; core protein; polyprotein
F;1-124/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 124;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 126; DB 2;
Pred. No. 2.5e-10;
5; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 126; DB 2
Pred. No. 8e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Data Library, December 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          84.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    84.68;
78.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: genomic RNA
A; Residues: 1-492 <SEE>
A; Cross-references: EMBL:X76918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similaricy
hes 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C;Accession: S41288
R;Seelig, R.
submitted to the EMBL Data L
A;Reference number: S41288
A;Accession: S41288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 22; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: mRNA
A; Residues: 1-411 <LIJ>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loc
Matches
```

tis C virus (strain U.K.) (fragment)
envelope protein El; envelope protein E2;

hepatitis

polyprotein

2

```
ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein - hepatitis C virus (genotype 1, N7) (fragment)
N;Contains: core protein
Species: hepatitis C virus
ariety: genotype 1, N7
Late: 19-May:1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                    #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Molecule type: genomic RNA
A;Residues: 1-115 <VAN>
A;Cross-references: EMBL:229452; NID:g443866; PIDN:CAA82590.1; PID:g443867
                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LIPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C;Accession: S41349
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPa.
A;Reference number: S41341
A;Accession: S41349
                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                             polyprotein - hepatitis C virus (genotype 1, N9) (fragment)
F;1-191/Product: core protein #status predicted <COR>
F;192-372/Product: envelope protein #status predicted <ENV>
F;373-492/Product: NS1 protein (fragment) #status predicted <NS1>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 115;
                                                                                              Length 492;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 115
                                                                                         Score 126; DB 2; Length 49
Pred. No. 9.5e-10;
5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 125, DB 2; Length 11
Pred. No. 3.2e-10;
3; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          , van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994 A; Description: Analysis of hepatitis C virus genotypes 1 to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A:Experimental source: genotype 1, N9
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capaid protein; core protein; polyprotein
F;1-115/Product: core protein *status predicted 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A:Experimental source: genotype 1, N7
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; polyprotein
F:1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 125; DB 2; I
Pred. No. 3.2e-10;
3; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                              N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 1, N9
Pate: 19-May 1994 #sequence_revision 26-Jul-1996
                                                                                                                                                                                                            1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                         Query Match
Best Local Similarity 78.6%;
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              83.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 Similarity 82.1%;
23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        larity 82.1%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
A; Cross-references: EMBL: 229454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
les 23; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                          .ccession: S41351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Accession: S41351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                     ö
                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
genome polyprotein - hepatitis C virus (strain JT)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nons
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Cispecies: hepatitis C virus
Cispecies: hepatitis C virus
Cispecies: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
Cispecies: 10: A45573
Ritanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijl Ritanaka, T.; Hijl Ritanaka, T.; Hijl Willes Res. 23, 39-53, 1992
A:Title: Molecular cloning of hepatitis C virus genome from a single Japanese carr A:Reference number: A45573; MUID:92295714; PMID:1318627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N'Contains: core protein
N'Contains: core protein
C; Species: hepatitis C virus
A,Variety: genotype 2, N5
A,Variety: genotype 2, N5
C; Accession: S41356
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
A; Reference number: S41341
A; Reference number: S41341
A; Reference number: S41341
A; Residues: 1-108 < VAN>A; Residu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (genotype 2, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 123; DB 2; I
Pred. No. 5.7e-10;
i; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A:Experimental source: genotype 2, N5
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           83.9%; Score 125; DB 2;
82.1%; Pred. No. 1.7e-09;
11ve 3; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     82.68;
78.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                larity 82.1%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
nes 23; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ഹ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

Gaps

ö

```
Nicontains: core protein:

C.Species: hepatitis C virus

A.Variety: genotype 2, N2

A.Variety: genotype 2, N2

A.Variety: genotype 2, N2

C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000

C.Accession: S41353

R.Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.

Submitted to the EMBL Data Library, January 1994

A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by Lippa.

A.Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein - hepatitis C virus (genotype 2, N2) (fragment)
                                                                                                                                                      DB 2; Length 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 108,
                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 121; DB 2; L
Pred. No. 1.1e-09;
4; Mismatches 2;
              A; Experimental source: genotype 3, N4 C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F; 1-123/Product: core protein #status predicted AMT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Cross-references: EMBL:229456
A;Experimental source: genotype 2, N2
C;Superfamily: hepatifis C virus genome polyprotein
C;Superfamils: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <MAI>
                                                                                                                                                   Score 122; DB 2;
Pred. No. 9e-10;
6; Mismatches
                                                                                                                                                                                                                                                                                         1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7, 2003, 11:21:49
                                                                                                                                                   81.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 81.2%;
Best Local Similarity 78.6%;
Matches 22; Conservative
                                                                                                                                             Ouery Match 81.9
Best Local Similarity 75.0
Matches 21; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: genomic RNA A; Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Search completed: August
Job time: 9.54545 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: S41353
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                  A Experimental Source: HCV-JT
A) Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; C; Seywords: ATP; glycoprotein; hydrolase; nucleotide CRC>
F; 2-115/Product: capsid protein C *status predicted CRC>
F; 109.729/Product: envelope protein Estatus predicted CMSI>
F; 300-729/Product: nonstructural protein NSI *status predicted cNSI>
F; 301-106/Product: nepacivirin #status predicted cNSI>
F; 301-1037/Region: nucleotide-binding motif B (P-loop)
F; 3112-1317/Region: nucleotide-binding motif B (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             submitted to the EMBL Data Library, April 1992
A:Description: A sentitive serodiagnosis of hepatitis C virus infection with two cloned
A:Reference number: S21336
                                                                                  Residues: 1-3010 <TAN>
Cross-references: GB:D11168; GB:D01171; NID:9221612; PIDN:BAA01943.1; PID:9221613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N,Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 3, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome polyprotein S29 (core protein region) - hepatitis C virus (fragment) C;Species: hepatitis C virus (C.Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 17-Nov-2000 C;Accession: S21336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Cross-references: EMBL: 229464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NID:959492; PIDN:CAA46517.1; PID:959493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
'014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (genotype 3, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 123; DB 1; Length su. Pred. No. 1.4e-08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rivan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994 A.Description: Analysis of hepatitis C virus genotypes 1 to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 122; DB 2;
Pred. No. 6.5e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                y Match 82.6%;
Local Similarity 78.6%;
hes 22; Conservative 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81.9%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A)Cross-references: EMBL:X65548;
C;Superfamily: hepatitis C virus
C;Keywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 81.99 est Local Similarity 78.69 atches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: genomic RNA
A; Residues: 1-88 <SAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Molecule type: genomic RNA Residues: 1-123 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Reference number: S41341
                                  Status: preliminary .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Accession: S21336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C; Accession: $41361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Juery Match
A;Accession:
A;Status: pre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
```

ö

Gaps

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec Run on:

US-09-491-146A-32 149 1 PKPKROTKRNTLRRPKNVKFPAGGQIVG 28 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

Searched:

127863 127863 seqs, 47026705 residues

:al number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

		Description	Q00269 h genome, po	hepatitis							h genome	c	ď	4	р депоше	5		P40267 lycopersico		_			_		Q9ugu0 homo sapien	Ogepd8 mus musculu	P47915 mus musculu	O27036 methanobact	Q9p0k1 homo sapien	_	volvox	000515 homo sapien		_	Q9kh11 thermus agu
SOMMAKIES		DI	POLG_HCVJT	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVTW	POLG_HCV1	PII1_MOUSE	POLG_BVDVN	H1_LYCPN	FURI_HUMAN	DYHC_DICDI	H12_RAT	RL29_YEAST	SGS4_DROME	DXR_ARATH	TF20_HUMAN	TF20_MOUSE	RL29_MOUSE	VATA_METTH	AD22_HUMAN	POLS_WEEV	H12_VOLCA		(7	RL5_RABIT	DTD_THEAQ
		DB	1	П	~	-				-	-	٦.		-	-	-	-	-	7	-	<b>,</b>		~		-	<b>.</b>		-	-	<del>, -1</del>	~	~		<b></b> 4 .	7
		Match Length	3010	513	220	520	737	737	3010	3010	3011	3033	3033	3010	3011	332	3988	202	794	4725	218	28	297	477	1960	1983	159	284	906	1236	240	517	3898	101	152
æ	Query	Match	2	Η.	∺	ä	Ξ.	;	ή.	81.2	81.2	ä	٠,	٠	75.8	4	4.	33.9	e.	'n	<del>.</del>		31.5	;	31.5	i.	ή.	31.2		;	ö	ö	ö	30.2	
		Score	123	N	$\sim$	121	$^{\circ}$	$\sim$	$\sim$	121	~	$\sim$	N,	114	113	52	51.5	50.5	20	49	47.5	47	47	47	47	47	46.5	٠.	ė.	•	46	4	45.5	45	45
	Result	Š.	-	7	m	4		9	7	80 (		10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	53	30	31	32	33

P02254 salmo trutt	P47962 mus musculu	P09895 rattus norv	P29673 drosophila	P23188 mus musculu	P21573 xenopus lae	Q04958 saccharomyc	Q9asu7 arabidopsīs	043405 homo sapien	P15180 saccharomyc	045882 clostridium	Q28193 bos taurus
H1_SALTR	RL5_MOUSE	RL5_RAT	APTE_DROME	FURI_MOUSE	YB1_XENLA	YMF9_YEAST	PPAN_ARATH	COCH_HUMAN	SYKC_YEAST	CGAA_CLOBI	FURI_BOVIN
ц,	-	Н	Н	<b>ب</b>	Н	-	~	-	Н	-	Н
194	296	296	469	793	303	1679	345	550	290	613	197
30.2	30.2	30.2	30.2	30.2	29.9	29.9	29.2	29.5	29.5	29.5	29.5
<b>4.</b>	4.5	45	45	45	44.5	44.5	44	44	44	44	44

### ALIGNMENTS

RESULT 1 1D POLGAROUS POLG
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope 41/voprotein El (GP32) (GP35); Envelope 91/voprotein E2
(GP68) (GP70) (NS1)] (Fragment).
Hepartits C virus (isolate HC-J2) (HCV).
Hepartits C virus (isolate HC-J2) (HCV).

NCBI\_TaxID-11111;

513 AA

STANDARD;

POLG\_HCVJ2 P27959;

PKPORKTKRNTYRRPQDVKFPGGGQIVG

```
POLG_HCVJ2
                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                          REMOVED FROM CAPSID PROTEIN C BY THE CELULULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN NSJ. FEZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ. FEZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ (POTENTIAL).
PROTEASE/HELICASE NSJ (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLIMERASE (POTENTIAL).
POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 123; DB 1; Length 3010;
Pred. No. 1e-09;
4; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC
                                                                                                             InterPro; IPR002868; HCV_NS5a.
InterPro; IPR002166; HCV_NS5a.
InterPro; IPR0071095; RNA_POI_DS_PS.
InterPro; IPR007094; RNA_POI_DS_PS.
InterPro; IPR007094; RNA_POI_DS_PS.
InterPro; IPR007094; RNA_POI_DS_PS.
InterPro; IPR007094; RNA_POI_PS_VII.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_NS1; 1.
Pfam; PF01050; HCV_NS2; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01054; HCV_NS5a; 1.
Pfam; PF01054; HCV_NS5a; 1.
Pfam; PF0057; Helicase_C; 1.
Pfam; PF0057; Helicase_C; 1.
Pfam; PF0057; Helicase_C; 1.
Pfam; PF006948; Viral_RARP; 1.
                              HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82.6%;
milarity 78.6%;
Conservative 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                326573
                                                       IPR002518; IPR004109; IPR000745; I
                                                                                                    IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ¥;
                                           IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity
Matches 22; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         623
645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                3D-structure.
INIT_MET
                                                       InterPro;
InterPro;
InterPro;
                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRANSMEN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation he Buropan Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                 ESCUENCE FROM N.A.

MEDLINE-92230232; PubMed=1314459;

MEDLINE-92230232; PubMed=1314459;

A Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,

A Tanaka T., Fukuda S., Tsuda F., Mishiro S.;

Tology to reported isolates: comparative study of four distinct genotypes.";

Uirology 188:331-341(1992).

LI FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

STAND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

STAND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

C. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN BAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CELLULAR AUMOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAGOR ENVELOPE PROTEIN (POTENTIAL).
MAGOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Interpro; IPR002522; HCV_capsid.
Interpro; IPR002521; HCV_core.
Interpro; IPR002519; HCV_core.
Interpro; IPR002519; HCV_env.
Interpro; IPR002531; HCV_MS1.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01559; HCV_mS1; 1.
Probom; PD186062; HCV_MS1; 1.
Probom; PD186062; HCV_MS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
Iransmembrane; Nonstructural REMOVED FROM CAPSID PROTEIN C BY THE INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D10074; BAA00968.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
```

ö

Gaps

; 0

```
Transmembrane;
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVHK
Q01403;
            NON_TER
SEQUENCE
                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                              POLG_HCVHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                   RESULT
            E S
                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR ANINOPERTIDAGE.

CAPSID PROTEIN C (POTENTIAL).

MAJOR ENVELOPE PROTEIN (POTENTIAL).

MAJOR ENVELOPE PROTEIN E (POTENTIAL).

NONSTRUCTURAL PROTEIN SI/E2 (POTENTIAL).

N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-93019030; PubMed-1383400;
MEDLINE-93019030; PubMed-1383400;
Abe K., Inchauspe G., Fujisawa K.;
Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2739(1992).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRRA.
                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66-OCT-2001 (Rel. 40, Last annotation update)
6600me polyprotein (Contains: Capsid protein C (Core protein) (P22);
670-009 (GP30) (RSI)] (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus (isolaté HCV-476) (HCV).
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   pfam; PF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Polyprotein: Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY
                                                          Score 121; DB 1; Length 513; Pred. No. 3.2e-10;
                                                                                                                       Indels
943F31E3514CDEF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                        520 AA
                                                                                                                    4; Mismatches
                                                                                                                                                                           1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                         InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521, HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PP01543; HCV_capsid; 1.
Pfam; PP01542; HCV_capsid; 1.
Pfam; PP01542; HCV_core; 1.
   55704 MW;
                                                          81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; D10688; BAA01530.1; -.
                                             Query Match
Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
   513 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1192
3384
1209
1209
1209
4214
4314
4314
4314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                               POLG_HCVH4
Q01404;
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
CHAIN
CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                     RESULT 3
POLG_HCVH4
ID POLG_H
AC Q01404
DT 01-JUL
DT 01-JUL
DT 16-OCT
CENTED CE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
S
                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME OF STREET O
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (FOTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Genome polyprotein (Contains: Capsid protein ( Core protein ( P22); Envelope glycoprotein E1 (GP32) (GP55); Envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).

Hepatitis C virus (isolate HCV-KF) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Abe K., Inchauspe G., Fujisawa K.,
"Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ( POTENTIAL) ( POT
                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                               Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR00251; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01550; HCV_env; 1.
ProDom; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                    Indels
     520
56499 MW; AA135246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CELLULAR AMINOPEPTIDASE
                                                                                               81.2%; Score 121; DB 1; Le
78.6%; Pred. No. 3.3e-10;
' wismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BY SIMILARITY.

N-LINKED (GLCNAC.

N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       520 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-93019030; PubMed-1383400;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; D10687; BAA01529.1; -. PIR; JQ1925; JQ1925.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                    Local Similarity 78.6 ies 22; Conservative
520 5
520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-31644;
```

```
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVJ7
P27961;
                         CARBOHYD
                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVJ7
                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ø
                                                                                                                                                                                                                                                                                                                                                                                                               Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HDD BRANK BRANK BD DE BERT BRANK BRANK BRANK BD BERT BRANK B
        g
                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.lsb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-70(-1992 (Rel. 23, Created)
01-70(-1992 (Rel. 23, Last sequence update)
16-00(-2001 (Rel. 40, Last annotation update)
16-00(-2001 (Rel. 40, Last annotation update)
66nome polyprotein (Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NSI and NS2] (Fragment).
Hepatitis C virus (isolate HC-J5) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HYDROPHOBIC: THE SMALL PROTEINS NS2B, NS2B, NS4A AND NS4B ARE HYDROPHOBIC: SUGGESTING A POSSIBLE MEMBRANE-ERLATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
LIPOPROTEIN THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-92230232; PubMed-1314459; Okamonto K., Lizuka H., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
  N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR00252; HCV_capsid.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_capsid; 1.
Pfam; PF01542; HCV_capsid; 1.
Pfam; PF01549; HCV_env; 1.
Pfam; PF01560; HCV_env; 1.
ProDom; PD186062; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                       Length 520;
                                                                                                                                                                 Indels
                                                                               MW; 1D2BD0A6FF27349B CRC64;
                                                                                                         81.2%; Score 121; DB 1; Le
78.6%; Pred. No. 3.3e-10;
Mismatches 2;
                                                                                                                                                                                                                                                                                                                                              737 AA
                                                                                                                                                                                                     1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                           32
                                                                                                                                                                                                                          EMBL; D10075; BAA00969.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genotypes.";
Virology 188:331-341(1992).
                                                                               56476
                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                              STANDARD;
424
431
449
520
520 AA;
                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11112;
                                                                                                                       Query ...
Best Local Similar
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Repactvirus
                                                                                                                                                                                                                                                                                                                                         POLG_HCVJ5
P27960;
                     CARBOHYD
CARBOHYD
                                                            NON_TER
SEQUENCE
  CARBOHYD
                                                                                                                                                                                                                                                                                                                      POLG_HCVJ5
ID POLG_HC
AC P27960,
DT 01-AUG
                                                                                                                                                                                                                                                                                                    RESULT
  FF FF S
                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                       g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66nome polyprotein (Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M3); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
Hepatitis C virus (isolate HC-J7) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-92230232; PubMed-1314459; Okada S.I., Yamamoto K., Lizuka H., Okada S.I., Vamamoto K., Lizuka H., Cokada S.I., Suraka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor monology to reported isolates: comparative study of four distinct
                                                                                                              POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 737;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 121; DB 1; Length 73
Pred. No. 4.7e-10;
4; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3AF699D82AD501B1 CRC64;
                                   N-LINKED (GLCNAC,
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPKROTKRNTLRRPKNVKFPAGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_MS1.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81207 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
347
2096
2099
2099
3005
4417
4430
4430
544
558
558
578
737
737
747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11114;
```

```
ö
                                                                                                       CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/F2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annoctation update)
16-Sep-2003 (Rel. 42, Last annoctation update)
16-Sep-2003 (Rel. 42, Last annoctation update)
Envelope polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GPV0) (NSI); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate BK) (HCV)
viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-91140698; PubMed-1847440; Takamizawa A., Moriakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; Structure and organization of the hepatitis C virus genome isolated
                                                                                             REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                 (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-96235224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81.2%; Score 121; DB 1; Length 737; 78.6%; Pred. No. 4.7e-10; 1ve 4; Mismatches 2; Indels
                                                               protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       737 AA; 81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                            N.LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                    POTENTIAL.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01559; HCV_core; 1.
Pfam; PF015560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Probom; Glycoprotein; Coat protein Glycoprotein; Coat protein Insamembrane; Nonstructural protein INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 78.6 les 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1487-1500.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON_TER
SEQUENCE
                                                                                                                                                                                                TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                       CHAIN
                                                                                                                                      CHAIN
                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 7
POLG_HCVBK
ID POLG_H
AC P26663
DT 01-AUG
                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loc
Matches
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.lsb-sib.ch/announce/or send an email to licensedlsb-sib.ch).
                                                                                                                                                                  WEDLINE-9827846; Pubmed-956891;

WEDLINE-9827846; Pubmed-956891;

WEDLINE-9827846; Pubmed-956891;

WA Yan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,

Steinkuehler C., Tomel L., de Francesco R., Kuo L.C., Chen Z.;

"Complex of NS3 protease and NS4A peptide of BK strain hepatitis C.

"Tomplex of NS3 protease and NS4A peptide of BK strain hepatitis C.

"Torein Sci. 7:837-847(1998).

"I FOWCION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE

"YDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

"STAND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

"C. - CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6

"DOSITION. Cys or Thi in Pil and Ser or Ala in Pi'.

"C. - CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                          X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE-97015088; PubMed-8861916;
Love R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
"The crystal structure of hepatitis C virus NS3 proteinase reveals a
trypsin-like fold and a structural zinc binding site.";
                                                                                                                                                                                                                                                                                                                                                                                                            SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN A AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; M58335; AAA72945.1; -. PIR; A38465; GNWVTC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDB, ICSJ, 08-NOV-99,
PDB, IGX5, 09-APR-02
PDB, IGX6, IO-APR-02,
PDB, IQUV, 26-JUN-00,
PDB, BOHM, 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1A10; 25-MAR-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF01006;
                                                                                                                                                                                                                                                                                                                                                                                              [RNA](N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PDB; 1NS3; (PDB; 1C2P;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     fam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    fam;
```

Gaps

ö

Score 121; DB 1; Length 3010; Pred. No. 2e-09; 4; Mismatches 2; Indels

28

us-09-491-146a-32.rsp

```
327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                     1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG
                                            Ouery Match 81.2%;
Best Local Similarity 78.6%;
Matches 22; Conservative
             1202
                          1688
                                                                                                                                                                                                                                                                          DISCUSSION OF SEQUENCE
1187 1188
1189 1197
1198 1202
1203 1204
1680 1688
3010 AA; 3
                                                                                                                                                                                                            NCBI_TaxID-11116;
                                                                                                                                                                                                       Hepacivirus
             HELIX
TURN
STRAND
SEQUENCE
 TURN
                                                                                                         POLG_HCVJA
                                                                                                    RESULT
                                                                                                                 51 FT ES
                                                                      ò
                                                                                셤
                                                            CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN (POTENTIAL).
MAJOR ENVELOR (POTENTIAL).
NONSTRUCTURAL PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEBAE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA DIRECTED RNA POLYMERASE (POTENTIAL).
                               Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                        REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                            (POTENTIAL)
                                                                                                                          POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                             (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                GLCNAC.
                                                                                                                                                                 N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                   GLCNAC
                                                                                                                                                                      N-LINKED
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SWART; SM00487; DEXDC; 1.
                                                                                      645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                           035
047
050
050
063
074
                                                                                                                                                                                                                                                                                                                                       085
087
094
097
097
103
                                                                                                                                                                                                                                                                                                                                                                                                     1133
                                                                                                                                                                                                                                                                                                                                       1082
1086
1090
1093
                                                                                                                                                                                                                                                                                                                                                                1095
                                                 3D-structure.
INIT_MET
                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                 CARBOHYD
                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                      CARBOHYD
                                                                                                                           TRANSMEN
                                                                                                                                                                                                                                                                                                       STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                          TURN
STRAND
STRAND
HELIX
TURN
TURN
STRAND
                                                                                                                                                                                                                                                                                                                          TURN
                                                                                                                                                                                                                                                                                                                                                   STRAND
                                                                                                                                                                                                                                                                                                                                                                STRAND
                                                                                                                                                                                                                                                                                                                                                                             TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                              STRAND
                                                                                                                                                                                                                                                                                            STRAND
                                                                                                                                                                                                                                                                                                                                                                                         STRAND
                                                                          CHAIN
CHAIN
CHAIN
                                                                                            CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                       HELIX
TURN
                                                                                                                                                                                                                                                                                                  HELIX
```

```
POLG_HCVJA STANDARD; PRT; 3010 AA.

POLG_HCVJA STANDARD; PRT; 3010 AA.

(C P26662;

01-AUG-1992 (Rel. 23, Last sequence update)

T 28-FEB-2003 (Rel. 41, Last annotation update)

T 28-FEB-2003 (Rel. 41, Last annotation update)

E Repeator of the contains: Capsid protein C (Core protein) (P22);

E Grenome polyprotein [Contains: Capsid protein C (Core protein) (P22);

E Grenome polyprotein [GR35]; Envelope glycoprotein E2 (GP69) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)

E C 34-22-); Protease/Helicase NS3 (P70) (Hepacivirin)

E C 34-22-); Protease/Helicase NS3 (P70) (Hepacivirin)

E NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].

E NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (RNA)(N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MINA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEGUENCE FROM N.A.
MEDINE-9108550; PubMed-2175903;
Kato N., Hijkata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Modecular cloning of the human hepatitis C virus genome from
"Appanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ.
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                 01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
Genome polyprotein (Contains capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Enveloge glycoprotein E2
(GP68) (GPV0) (NSI); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.29.9); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Monstructural protein NS5A (P65); Nonstructural protein NS5A (P65); Nonstructural protein NS5A (P65); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate H) (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187644;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-TAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-98154321; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatlis C virus NS3 RNA helicase domain with a bound
oligonucleotide: the crystal structure provides insights into the mode
of unwinding.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -i- FUNCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-1- FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS4A, NS4A-NS4B, NS4B-NS5A, NS4B-NS5A, NS4B-NS5B, NS4B-NS5B, NS4B-NSAB, NS4A-FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE ACTIVATION OF NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.

CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (RNA)(N).
--- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MINA.
--- PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
--- SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39.
--- SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FUNCTION: NSSA SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE. FUNCTION: NSSB IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN ESSENTIAL ROLE IN THE VIRUS REPLICATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Prince A.M.; "Genomic strain H of hepatitis "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                         PKPKRQTKRNTLRRPKNVKFPAGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Structure 6:89-100(1998)
                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-11108;
                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus.
                                                                                                                                     POLG_HCVH
                      -
                                                     S
                                                                                                                       POLG_HCVH
                                                                                                      RESULT
                                                   g
                                                                                                                                       CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 121; DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED GLCNAC.
                                          InterPro; IPR001410; DEAD.
InterPro; IPR001522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR001409; HCV_NS3.
InterPro; IPR001499; HCV_NS3.
InterPro; IPR001499; HCV_NS3.
InterPro; IPR001496; HCV_NS3.
InterPro; IPR001650; HCV_NS3.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001650; HCV_Capsid.
InterPro; IPR001650; Helicase_C.
InterPro; IPR01650; HCV_Capsid.
InterPro; IPR001650; HCV_Capsid.
InterPro; IPR00160; HCV_NS1; I.
IPfam; PF01501; HCV_NS2; I.
IPfam; PF01001; HCV_NS3; I.
IPfam; PF01001; HCV_NS4; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2077
2240
2788
327017 M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00487; DEXDC;
               MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2041
2077
2240
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
```

Gaps

ö

2; Indels

Pred. No. 2e-09;

78.68;

Conservative

Best Local Similarity Matches 22; Conserv

O

ö Gaps ö

AREAN A STANKER A STANKER

```
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P60) (P70) (RNA-directed RNA polymerase; (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-38) (HCV).
                                                                                                                                                                                                                                                                                                                                        SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                      SIMILARITY)
SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 121; DB 1; Length Ju-Pred, No. 2 le-09; 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CRC64;
                                                                                                                                                                                                                                                                                                                                      (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1091 N-LINKED (GLCNAC. . . 2038 N-LINKED (GLCNAC. . . 2811 N-LINKED (GLCNAC. . . . AA; 329165 MW; F957F5C1A273HE9E
                                                                                                                                                                                                                                                                                                                                   CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; heltcase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SWART; SW00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                   1010
1619
1866
2017
3033
369
1087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus.
NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     649
1091
2038
2811
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVJ8
P26661;
                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                   PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                     A MEDLINE-9204440: PubMed-1658196;
A MEDLINE-9204440: PubMed-1658196;
A MEDLINE-9204440: PubMed-1658196;
A Okamoto H., Okada S.-I., Suglyama Y., Kurai K., Lizuka H.,
Machida A., Miyakawa Y., Mayumi M.;
Mucleotide sequence of the genomic RNA of hepatitis C virus isolated
from a human carrier: comparison with reported isolates for conserved
from a human carrier: comparison with reported isolates for conserved
and divergent regions.
Tyle J. 22.2697-2704(1991).
J. Gen. Virol. 72.2697-2704(1991).
J. Gen. Virol. 72.2697-2704(1991).
C -1- FUNCTION: THE SMALL PROFIEIDS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in Pl and Ser or Ala in Pl'
C -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
          01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Bnvelope glycoprotein El (G922) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22.-); Proteases/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS5A (P4); Nonstructural protein NS5A (P65); Nonstructural protein NS5A (P66); P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-J6) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (RNA)(N).
SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROPEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pro: IPR007095; RNA_pol_DS_PS.
Pro: IPR007094; RNA_pol_PSvir.
PF01543; HCV_capsid; 1.
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_NS5a.
HCV_RdRP.
Helicase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR00252; HCV_capsid.
IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR002519; HCV_env.
IPR002531; HCV_NSI.
IPR002518; HCV_NS2.
IPR004109; HCV_NS3.
IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D00944; BAA00792.1; -. PIR; JQ1303; JQ1303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_NS4a; 1
HCV_NS4b; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_NS3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR001490;
IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                               NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01560;
PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PF01006;
PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
```

÷

ö

Gaps

ö

MEROPS;

Pfam; Pfam; Pfam; Pfam; Pfam;

```
01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 42, Last aenoctation update)
16-SEP-2003 (Rel. 42, Last annotation E2
16-SEP-2003 (Rel.) Frotein Protein Protein Protein NSA (Peparitin)
16-SEP-2003 (Rel.) Frotease/hellcase NS3 (P70) (Reparitin)
16-SEP-2003 (Rel. 42, Lel. 42, Lel. 42, Lel. 43, Rel. 44, Lel. 44, Lel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 121; DB 1; Length 3033;
Pred. No. 2.1e-09;
4; Mismatches 2; Indels (
                                                                (BY
(BY
                                                           CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                         N-LINED GLCNAC
                                                                                                                                         ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. 2.16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        330177 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLG_HCVTW
P29846;
        CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A PART TARY NA PAR
           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VICTORY 188:331-341(1992).

-I- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-I- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

-I- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
SEQUENCE FROM N.A.
MEDLINE-92230232; Pubmed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.
Fukuda S., Tsuda F., Mishiro S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                           "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_core.
HCV_core.
HCV_env.
HCV_NSI.
HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF01538; HCV_NS2; 1.
Pfam; PF012907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01506; HCV_NS4s; 1.
Pfam; PF01506; HCV_NS5s; 1.
Pfam; PF01968; VITAL_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; D10988; BAA01761.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001410;
InterPro; IPR00252;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002519;
InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR; A40250; GNWVJ8.
HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Core protein; C
Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                               (RNA)(N).
                                                                                                                                                                genotypes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS
```

ö

Gaps

```
RESULT 13
POLG_HCV1
                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropaen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CORE PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN NSI/RZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSZ (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FIGURE, ALTECTOR DEXDC: 1.
SMART; SMO0487; DEXDC: 1.
POlyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHARGE RELAY SYSTEM (BY S. CHARGE RELAY SYSTEM (BY S. CHARGE RELAY SYSTEM (BY S. ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_Env.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR000409; HCV_NS2.
InterPro; IPR0004409; HCV_NS4.
InterPro; IPR0004409; HCV_NS4b.
InterPro; IPR000460; HCV_RS4b.
InterPro; IPR007095; RNA_POl_DS_PS.
InterPro; IPR007095; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_PSvir.
Pfam; PF01543; HCV_cappid; 1.
Pfam; PF01543; HCV_cappid; 1.
Pfam; PF01543; HCV_Cappid; 1.
Pfam; PF01543; HCV_NS1; 1.
Pfam; PF01539; HCV_NS1; 1.
Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4b; 1.
                                                                                                                                                                                                                                                                            EMBL; M84754; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       115

1283

1811

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006

10006
                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro, IPR001410,
InterPro, IPR002522,
InterPro, IPR002521;
                                                                                                                                                                                                                                                                                                             PIR; A40244; GNWVTW.
PDB; 1N64; 25-FEB-03
PDB; 1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
  CARKER KARE AND DRAWKEN WAS A STATE OF THE S
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE—91172826; PubMed=1848704;

MEDLINE—91172826; PubMed=1848704;

MEDLINE—91172826; PubMed=1848704;

MEDLINE—91172826; PubMed=1848704;

RA Gallegos C., Coit D., Medina—Selby A., Barr P.J., Weiner A.J.,

RA Gallegos C., Coit D., Medina—Selby A., Barr P.J., Weiner A.J.,

RA Bradley D.W., Kuo G., Houghton M.;

RA Bradley D.W., Kuo G., Houghton M.;

RA Bradley D.W., Mon G., Houghton M.;

ROC: NEAL Acad. Sci. U.S.A. 88:2451-2455(1991).

PICC. NEALL PROTEINS NS2B, NS4B AND NS4B ARE

HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE—RELATED FUNTION.

C. I- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6

DOSITION, Cys or Thr in P1 and Ser or Ala in P1.

C. I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseé(isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
66nome polyprotein (Contains: Capsid protein (Core protein) (P22);
66nome polyprotein (Contains: Capsid protein (Core protein E2)
6700 (NS1); Protein P7; Nonstructural protein NS2 (P21)
673.4.22...); Protease/Helicase NS3 (P70) (Hepacivirin)
673.4.22...); Protease/Helicase NS3 (P70) (Hepacivirin)
673.4.22...); Nonstructural protein NSAA (P4); Nonstructural protein NSAB (P60); Nonstructural protein NSAA (P60); Nonstructural protein NSAB (P60); Nonstructural protein NSAB (P60); Nonstructural protein NSAB (P70); Nonstructural pr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
          (POTENTIAL)
                                               POTENTIAL
                                                                          (POTENTIAL)
                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 114; DB 1; Length 3010;
Pred. No. 2.2e-08;
4; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC. .) (POTI
MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                    (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 KPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                76.5%;
77.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                               327047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 76.5
Best Local Similarity 77.8
Matches 21; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                              623
645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                  Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11104;
                                                                                                                                                                                              623
645
2041
2077
2240
2529
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (RNA)(N)
CARBOHYD
CARBOHYD
                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLG_HCV1
                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P26664
```

Gaps

```
EXCLUENCE FROW N.A.

RX AMAIN—21085660; PubMed=11217851;

RA AMAIN—21085660; PubMed=11217851;

RA ATARWA T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

A ATARWA T., Hara M., Nishi K., Konno H., Adachi J., Fukuda S.,

A Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Alzawa K., Izawa M., Ashburner M., Balalov S., Casavant T.,

RA Alzawa K., Matsud H.A., Ashburner M., Balalov S., Casavant T.,

RA Eleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

Ruchl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

Rohriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washlo T.,

Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washlo T.,

Schriml L.M., Staubli F., Suzuki R., Tomita M., Gariboldi M.,

Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

A Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Willaling L.,

Mynshaw Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,

Hayashizaki Y.,

Hayashizaki Y.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Liao C., Zhao M., Li T.;
"The expression of mouse LPTS1, a homolog of human tumor suppressor
LPTS, in mouse liver.";
Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                  Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                                 PIII_MOUSE STANDARD; PRT; 332 AA.
09C2X5; Q9IW29; Q9D0C2;
28-FEB-2003 (Rel. 41, Last sequence update)
Pin2-interacting protein X1 (TRF1-interacting protein 1) (Liver-related putative tumor suppressor) (LPTS1) (67-11-3 protein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'Functional annotation of a full-length mouse cDNA collection.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE-21558139; PubMed-11701125;
Zhou X.Z., Lu K.P.,
"The Pin2/TRF1-interacting protein PinX1 is a potent telomerase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 113; DB 1; Length 3011;
Pred. No. 3.1e-08;
    327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                                                      5; Mismatches
                                                                                                                              1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                   / Match 75.8%;
Local Similarity 71.4%;
les 20; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 38-332 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 speckles (By similarity)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nature 409:685-690(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cell 107:347-359(2001).
    3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IISSUE-Embryo;
                                                                                                                                                                                                                                                                                                                                                                                                                 PINXI OR LPTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-ICR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       inhibitor
      SEQUENCE
                                              Query Match
                                                                     Best Loc
Matches
                                                                                                                                                                                                                               RESULT 14
                                                                                                                                                                g
                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN SI, FEZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1, FEZ (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
CHARGE RELAY SYSTEM (BY S.
CHARGE RELAY SYSTEM (BY S.
CHARGE RELAY SYSTEM (BY S.
ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                  IPR002866; HCV_NSSA.
IPR002166; HCV_RGRP.
IPR001650; Hellcase_C.
IPR007095; RNA_pol_DS_PS.
                                                                                                                          HCV_core.
                                                                                                                                                                                                                                                                                                                                                                Interpro: IPR007094; RNA_DOL_PS: Fdam; PP01543; HCV_capsid; 1. Pfam; PP01542; HCV_capsid; 1. Pfam; PP01542; HCV_capsid; 1. Pfam; PP01569; HCV_RN2; 1. Pfam; PP01560; HCV_RN3; 1. Pfam; PP010507; HCV_RN3; 1. Pfam; PP01001; HCV_RN3; 1. Pfam; PP01001; HCV_RN3; 1. Pfam; PP01001; HCV_RN3; 1. Pfam; PP01051; HCV_RN3; 1. Pfam; PP01506; HCV_RN3; 1. Pfam; PP01506; HCV_RN3; 1. ProDom; PD186062; HCV_RN3; 1. ProDom; PD186062; HCV_RN3; 1. ProDom; PD186062; HCV_RN3; 1. PRODOM; DEXDO; 1.
                                                                                                                                                                                                                                                HCV_NS4a.
                                                                                                                        InterPro; IPR002522; H
InterPro; IPR002521; H
InterPro; IPR002519; H
InterPro; IPR002531; H
PIR; A39166; GNWVC3.
PDB; 1A1V; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
                                                                                                                                                                                                                         IPR004109;
IPR000745;
IPR001490;
                                                                                                                                                                                                                                                                                                                            InterPro; IPR001650;
InterPro; IPR007095;
                                                                                                         InterPro; IPR001410;
                                                             MEROPS; $29.001;
                                                                                  U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3D-structure.
INIT_MET
                                                                                                                                                                                                        InterPro; InterP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                  MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
```

```
Search completed: August 7, 2003, 11:20:05
Job time : 5.90909 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
   Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-8826585; Pubmed-2838958;
MEDLINE-8826585; Pubmed-2838958;
Collett M.S., Larson R., Belzer S.K., Retzel E.;
Collett M.S., Larson R., Belzer S.K., Retzel E.;
Collett M.S., Larson R., Belzer S.K., Retzel E.;
Virology 165:200-208(1988).
-1- FUNCTION: PESTIVIRUS P80 (P125) MAY BE A BIFUNCTIONAL PROTEIN WITH HELICASE AND PROTEASE ACTIVITY.
-1- PUM: COLL GIVES RISE TO GP62 AND GP53; GP62 IN TURN YIELDS GP48
                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-88225858;
Collett M.S., Larson R., Gold C., Strick D., Anderson D.K.,
Purchio A.F.;
"Molecular cloning and nucleotide sequence of the pestivirus bovine
viral diarrhea virus.",
Virology 165:191-199(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Virusės; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bovine viral diarrhea virus (isolate NADL) (BVDV) (Mucosal disease
                                                                                                                                                                                                                    GO; GO:0005730; C:nucleolus; IDA.
GO; GO:0005730; C:nucleolus; IDA.
GO; GO:0005515; F:protein binding activity; IPI.
GO; GO:0007004; P:telomerase-dependent telomere maintenance; IDA.
InterPro; IPRO10467; G_Patch.
Pfam; PF01585; G-patch; 1.
PROSITE; PS50174; G_PATCH; 1.
Anti-onogenes; Nuclear protein; Chromosomal protein; Telomere.
DOMAIN
                                                                                                                                                                                                                                                                                                                                                  TELOMERASE INHIBITORY DOMAIN (TID)
                                                                                                                                                                                                                                                                                                                                                                                                                           ö
 DOMAIN: The TID (telomerase inhibiting domain) domain is sufficient to bind TERT and inhibits its activity (By
                                                                                                                                                                                                                                                                                                                                                                                                 DB 1; Length 332;
2.9;
                                                                                                                                                                                                                                                                                                                                                                                                                          7; Indels
                                                                                                                                                                                                                                                                                                                                                             ETPV -> HAPC (IN REF. 3).
6B7147CED58D991A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-FEB-1991 (Rel. 17, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
01-FEB-1996 (Rel. 33, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 3988 AA
                                                                                                                                                                                                                                                                                                                                                                                                34.9%; Score 52; DB 40.9%; Pred. No. 2.9; ive 6; Mismatches
                                                                                                                                             EMBL; AF421879; AAL32445.1; -.
EMBL; AF321817; AAL37221.1; -.
EMBL; AK012057; -: NOT_ANNOTATED_CDS.
EMBL; AK011578; -: NOT_ANNOTATED_CDS.
EMBL; AJ344106; CAC51439.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              |||||: : |:||: |
PKPKKRKAKKKLQRPEGVEIDA 316
                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPA 22
                                                                                                                                                                                                                                                                                                                                                             210 213 E'
332 AA; 37221 MW;
                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match
Best Local Similarity 40...
9; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                             MGI:1919650; Pinxl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11100;
                        similarity).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_BVDVN
P19711:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pestivirus
                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          virus).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 15
POLG_BVDVN
ID POLG_BV
AC P19711;
DT 01-FEB-
DT 01-FEB-
                                                                                                                                                                                                              MGD;
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein; Glycoprotein; Helicase; Serine protease; Hydrolase. CHAIN 1 7270 P20 (30KD). CHAIN 727 7 7 GPI15/GP62-GP53 (GLYCOPROTEIN). CHAIN 7 GPI25/GP54-GP80.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 1; Length 3988;
-i- SIMILARITY: TO THE HOG CHOLERA VIRUS GENOME POLYPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4474212F338661BB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. 47;
5; Mismatches
                                                                                                                                                                                                                                                            InterPro; IPR001280; CDV1r_endptseP80.
InterPro; IPR001410; DEAD.
InterPro; IPR001410; DEAD.
InterPro; IPR00165; HCV_RGRP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001059; RNA_POl_DS_PS.
InterPro; IPR001569; RNA_POl_DS_PS.
InterPro; IPR001569; RNASE_T2.
Pfam; PF00271; helicase_C; I.
Pfam; PF00998; Viral_RGRP; I.
PRNNS; PR00729; CDVBNDOPTASE.
SMART; SM00487; DEACC; I.
SMART; SM00490; HELICC; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 51.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3189 PKPKQVTKEAVRNLIEQKKDVEIP 3212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPKRQTK---RNTLRRPKNVKFP
                                                                                                                                                                                 EMBL; M31182; AAA42854.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              34.6%;
45.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             449154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best_Local Similarity 45.8
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ä
                                                                                                                                                                                                 PIR; A29198; GNWVBV
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                MEROPS; C53.001; -. MEROPS; S31.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1451
1803
22224
23307
2772
2772
2981
3778
3867
3988
```

us-09-491-146a-32.rspt

```
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
                                                                                                                                                                                                          hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                    hepatitis
hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                           hepatitis
hepatitis
hepatitis
hepatitis
                                                                                                                                                          hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               hepatitis
hepatitis
                                                                                                                                                                                               hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Chayama K.;
Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL: D49374; BAA08372.1; --
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
Hepatitis C virus type 3b.
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
081233
081231
081234
081234
081234
068861
068863
068863
068863
068951
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
068308
06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3023 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ALIGNMENTS
081233
0812331
0812342
0812342
068863
068863
068863
068863
068863
068308
068308
068308
068308
068308
068308
068308
                                                                                                                                                                                                                                                                                                                                                                       080P72
080P71
080P74
080P80
                                                                                                                                                                                                                                                                                                                                                                                                                                                             080P84
080P75
080P79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pro; IPR007095; RNA_pol_DS_PS.
Pro; IPR007094; RNA_pol_PSvir.
PF01543; HCV_capsid; 1.
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS4a.
HCV_NS5a.
HCV_RGRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Interpro; IPR002519; H
Interpro; IPR002531; H
Interpro; IPR004109; H
Interpro; IPR004109; H
Interpro; IPR00145; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
NCBI_TaxID-42791;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROPS; S29.001;
MEROPS; U39.001;
nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          081487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             081487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 1
081487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hepatitis of hepat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    hepatitis (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q81487 hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   hepatitis
                                                                                                                August 7, 2003, 11:05:41; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   is the number of results predicted by chance to have a
nter than or equal to the score of the result being printed,
ived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   081550
081772
084782
091725
091725
091726
068413
068305
08126
081230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P87760
P87761
                                                                                                                                                                                                                                                                                                                                                    830525
                               Compugen Ltd.
           GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compug
                                                                                                                                                                                                                                                                                                                                               tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                        149
1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                               830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                  protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                P87760
P87761
Q8JYR5
Q8JYR2
Q8DYR2
Q9JK25
Q9JK27
Q9JK27
Q9JK27
Q9JK27
Q6B413
Q8J28
Q8J23
Q8J233
                                                                                                                                                                                                                                                            BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_fungi:*
sp_human:*
sp_invertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     score greater than or equal to
and is derived by analysis of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_bacteriap: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                   seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_archea:*
sp_bacteria:*
                                                                                                                                                                                        US-09-491-146A-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_rvirus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_plant:*
sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  419
108
108
415
415
100
191
191
191
43
43
109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          87.9
87.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No.
                                                                                                                                                                                                          Perfect score:
                                                                                                                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  132
132
132
133
133
130
130
130
                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB s
Maximum DB s
                                                                                  OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Database :
                                                                                                                                                                                                                            Sequence:
                                                                                                                                                                                                                                                                                                               Searched:
                                                                                                                     Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ş
```

Gaps

ö

```
Pfam; PF01560; HCV_NS1; 1.

PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.

Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.

CHAIN 1 >190
                                                                                                                                                                                                                                                                                                                P87760 PRELIMINARY: PRT: 108 AA.
P87756;
01-MAY-1997 (TEMBLrel. 03, Created)
01-MAY-1997 (TEMBLrel. 03, Last sequence update)
01-OCT-2002 (TEMBLrel. 22, Last annotation update)
01-OCT-2002 (TEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-96305314; PubMed-8712927; Ohloo E., Ohba K.-I., Wu R.-R., Ohno T., Mizokami M., Saleh M.G., Orito E., Ohba K.-I., Wu R.-R., Kolde T., Tibbs C.J., Nouri-Aria K.T., Tokudome S., Williams R.; "Usefulness and limitation of phylogenetic analysis for hepetitis C virus core region: application to isolates from Egiptian and Yemeni
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ohno T., Mizokami M.;
"Determination of nine genotypes of hepatitis C virus using PCR method.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 88.6%; Score 132; DB 12; Length 108; Best Local Similarity 85.7%; Pred. No. 2.1e-12; Matches 24; Conservative 2; Mismatches 2; Indels (
                                                                                                                                          Length 419;
                                                                                                                                                                            1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein.
NON_TER 108 108
SEQUENCE 108 AA; 11773 MW; A05A3836EC14BA21 CRC64;
                                                                             191 >419 ENVELOPE PROTEIN.
419 419
419 AA; 45391 MW; CB8956E32EA6DD6C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Thesis (1995), Nagoya City University Medical School
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P87761 PRELIMINARY; PRT; 108 AA. P87761; 108 AA. P87761; 01-MAY.1997 (TrEMBLrel. 03, Created) 01-MAY-1997 (TrEMBLrel. 03, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) Core protein (Genome polyprotein) (Fragment).
                                                                                                                                         Score 137; DB 12;
Pred. No. 1.6e-12;
2; Mismatches 1;
                                                                                                                                                                                                        1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                          5 PKPQRQTKRNTLRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPQRQTKRNTTRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                          Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                             CHAIN
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                    RESULT 3
P87760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P87761
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    **Numedotide sequence hepatitis C virus (type 3b) isolated from a Juncaleotide sequence hepatitis C virus (type 3b) isolated from a Japanese patient with chronic hepatitic C.";

J. Gen. Virol. 75:3823-3828(1994).

C -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN A MAD GAITCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF REABL; DI1443; BAA02007.1;

R InterPro; IPRO02521; HCV_capsid.

R InterPro; IPRO02521; HCV_core.

R InterPro; IPRO02531; HCV_core.

R InterPro; IPRO02531; HCV_core.

R InterPro; IPRO02531; HCV_core.

R InterPro; IPRO02531; HCV_core.

R InterPro; IPRO02543; HCV_core.

R Pfam; PFO1543; HCV_core: 1.
                                                                                                                            ProDom; PD186062; HCV_NS1; 1.
SWART; SMOU487; DEXDC: 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
CHAIN 191 81 E1.
CHAIN 192 884 E1.
CHAIN 385 737 E2/NS1.
                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
MEDLINE-94270990; PubMed-7545932;
Hotta H., Doi H., Hayashi T., Purwanta M., Soemarto W., Mizokami M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Analysis of the core and El envelope region sequences of a novel variant of hepatitis C virus obtained in Indonesia.";
Arch. Virol. 136:53-62(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-95088611; Pubmed-7996156;
Chayama K., Tsubota A., Koida I., Arase Y., Saitoh S., Ikeda K.,
                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                100.0%; Score 149; DB 12; Length 3023; 100.0%; Pred. No. 2e-13; 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                                                                                                                                    329734 MW; 5268D1EC410AC545 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              419 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Genome polyprotein (Fragment).
Hepatitis C virus.
                                                              Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4D; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01542; HCV_core; 1
PF01539; HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                        1014
1665
1980
3023
                                                                                                                                                                                                                                       192 31
385 77
738 101
1015 166
1666 191
1981 3023 AA;
                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ohba K., Homma M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11103;
                                  PF01538; PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                                                                                                                                                                                                                   CHAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            081345
                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam:
    Pfam;
                                                   Pfam;
                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                 Matches
   DR DR DR DR DR KW KW KW KW LFT FFT FFT FFT FFT FFT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
```

ö

Gaps

```
EMBL; D16616; BAA04038.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-RIG304;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
SEQUENCE
                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                             OBJYR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q8QP82
                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
08JYR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      080P82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AC OC OC OC OC
 STEERS
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Tokita H., Shrestha S.M., Okamoto H., Sakamoto M., Horikita M.,
Tokita H., Shrestha S.M., Okamoto H., Mayumi M.;
Ilzuka H., Shrestha S., Miyakawa Y., Mayumi M.;
Hepatitis C. virus variants from Nepal with novel genotypes and their
classification into the third major group.";
J. Gen. Virol. 75:931-936(1994).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                    BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                        Ohno T., Mizokami M., Saleh M.G., Orito E., Ohba K.-I., Wu R.-R., Kolde T., Tibbs C.J., Nouri-Aria K.T., Tokudome S., Williams R., "Usefulness and limitation of phylogenetic analysis for hepatitis C virus core region: application to isolates from Egiptian and Yemeni
     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                          FACH. VICOL. 141:1101-1113(1996).

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, D82033; BAA11518.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                   Ohno T., Mizokami M.; "Determination of nine genotypes of hepatitis C virus using PCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Shrestha S.M., Tsuda F., Okamoto H., Tokita H., Horikita M., Tanaka T., Miyakawa Y., Mayumi M.; "Hepatitis B virus subtypes and hepatitis C virus genotypes in patients with chronic liver disease in Nepal."; Hepatology 19:805-809(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Core, env, and part of E2/NS1 (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11773 MW; A05A3836EC14BA21 CRC64;
                                                                                                                                                      Thesis (1995), Nagoya City University Medical School
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (TrEMBLrel. 01, Last sequence update) (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 132; DB 12;
Pred. No. 2.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            415 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRQTKRNTTRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                         MEDLINE-96305314; PubMed-8712927;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-NE137;
MEDLINE-94201770; PubMed-8151307;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-94186155; PubMed-8138250;
                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002522; HCV_caps1d. Pfam; PF01543; HCV_caps1d; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sch 88.6%;
11 Similarity 85.7%;
24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1996 (TrEMBLrel. 01-OCT-2002 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            108 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                  SEQUENCE FROM N.A.
Hepatitis C virus
                                               NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-NE137
                                                                                                                                                                                                                                                                                                   patients.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          081550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 5
081550
ID 08151
AC 08151
DT 01-N
DT 01-N
DE COTE
OS HEPA
```

õ g

SOFTE

```
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_core.
Pfam; PR01543; HCV_capsid; 1.
Pfam; PR01543; HCV_capsid; 1.
Pfam; PR01559; HCV_core; 1.
Pfam; PR01559; HCV_core; 1.
Pfam; PR01560; HCV_NS1; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein; Polyprotein; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepartitis C virus in Western Siberia.";
"Genetic variability of hepartitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-!- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN WAND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                               Length 415;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 73;
                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                          OAD63410FA8F522F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            73 AA; 8272 MW; 64D60827A48FD796 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-0CT-2002 (TrEMBLrel, 22, Created)
01-0CT-2002 (TrEMBLrel, 22, Last sequence update)
01-MAR-2003 (TrEMBLrel, 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           87.9%; Score 131; DB 12;
82.1%; Pred. No. 1.9e-12;
iive 4; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                          Score 132; DB 12;
Pred. No. 9.1e-12;
                                                                                                                                                                                                                                                                                                                       88.6%; Scor.
85.7%; Pred. No. y...
... 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            73 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTIRRPQNVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AF506623; AAM33399.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                          415 AA; 44703 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match
Best Local Similarity 82.10,
10cal 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                  Local Similarity 85.7 nes 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus.
NCBI_TaxID=11103;
```

ö

```
MEDLINE-20273994; PubMed-10811932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                87.9%;
ilarity 82.1%;
Conservative 4
     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ouery Match
Best Local Similarity
Matches 23; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                STRAIN-SOM1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-SOM2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091K26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         091K26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 10
Q91K26
     g
                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                         Kalinina O., Norder H., Mukomolov S., Magnius L.O.;

"A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg.";

J. Virol. 76:4034-403(2002).

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PRADI. AV070183; AALS8597.1;

InterPro: IPR002522; HCV_capsid.

Petan: PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-2003;
MEDLINE-2073994; PubMed-10811932;
Abid K., Quadri R., Veuthey A.L., Hadengue A., Negro F.;
And K., Quadri R., Veuthey A.L., Hadengue A., Negro F.;
A novel hepatitis C virus (HCV) subtype from Somalia and its
classification in HCV clade 3.";
J. Gen. VIROL 81.1485-1493 (2000)
I. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                           uery Match 87.9%; Score 131; DB 12; Length 100; est Local Similarity 82.1%; Pred. No. 2.7e-12; Matches 23; Conservative 4; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87.9%; Score 131; DB 12; Length 191;
82.1%; Pred. No. 5.5e-12;
1ve 4; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
Core structural protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                      Polyprotein.
NON_TER 100 100
SEQUENCE 100 AA; 11235 MW; F2A0FD95F9E85AEE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein.
NON_TER 191 191
SEQUENCE 191 AA; 20730 MW; 7E85DDDE0DC4AC3A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (TrEMBLrel. 15, Created)
(TrEMBLrel. 15, Last sequ
(TrEMBLrel. 22, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                         MEDLINE-21904745; PubMed-11907242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; AF216794; AAF44741.1; -.
InterPro; IPR002522; HCV_copsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_copsid.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 82.1
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                     FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11103;
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              091K25;
01-OCT-2000 (
01-OCT-2000 (
```

**091K25** 

RESULT 8 **Q91K25** 

ð a Query Match

ò a

```
Abid K., Quadri R., Vetthey A.L., Hadengue A., Negro F.;

"A novel hepatitis C virus (HCV) subtype from Somalia and its
"T classification in HCV clade 3.";
"J. Gen. Virol. 81:1485-1493(2000).

"I. Guburt: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN BYNELOPE CONSISTS OF TWO PROTEINS:

"EMBL: AF216792; AAF44739-1; ---
"R InterPro; IPR002522; HCV_capsid.
"R InterPro; IPR002521; HCV_capsid.
"R Pfam; PF01542; HCV_capsid; I.
"R Pfam; PF01542; HCV_capsid; I.
"R Pfam; PF01542; HCV_capsid; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Abid K., Quadri R., Veuthey A.E., Hadengue A., Negro F.;
Abid K., Quadri R., Veuthey A.E., Hadengue A., Negro F.;
A novel hepatitis C virus (HCV) subtype from Somalia and its
classification in HCV clade 3.";
J. Gen. Virol. 81:1465-1493(2000)
-1- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; AF216793; AAF44740.1;
                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 191;
                                    01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core structural protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-0CT-2002 (TrEMBLrel. 22, Last annotation update)
Core structural protein (Genome polyprotein) (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               191 191
191 AA; 20730 MW; 7E85DDDE0DC4AC3A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20730 MW; FC844F8E1573F615 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 131; DB 12;
Pred. No. 5.5e-12;
4; Mismatches 1;
191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPORKTKRNTIRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein.
NON_TER 1
SEQUENCE 19
```

Gaps

ö

2; Indels

Length 415;

```
Pfam; PF01539; HCV_env; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.
NON_TER 415 415
SEQUENCE 415 AA; 44881 MW; FD151EF15R1C:1EC CEC61.
-i- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                         Score 131; DB 12;
Pred. No. 1.3e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPORQTKRNTPRRPRNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                       InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002531; HCV_car.
InterPro; IPR002531; HCV_car.
Pfam; PF01543; HCV_capsid; I.
Pfam; PF01542; HCV_capsid; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       87.2%;
illarity 82.1%;
Conservative 4
                                                                                                                                                                                                                                                                                                                                                                                     Match 87.9%;
Local Similarity 85.7%;
les 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22,
23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q8JYR3 PRELIMINARY;
Q8JYR3;
01-OCT-2002 (TTEMBLTE1. 2
01-CCT-2002 (TTEMBLEE1. 2
01-MAR-2003 (TTEMBLEE1. 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
es 23; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-HCV-BB46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         968305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 14
Q8JYR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 13
Q68305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    44441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SO DR REPRESENTATION OF THE SO DE TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     five novel
9c) major
                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-97052554; PubMed-8897188; Bernier L., Willems B., Delage G., Murphy D.G.; Identification of numerous hepatitis C virus genotypes in Montreal, Canada.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             J. Clin. Microbiol. 34:2815-2818(1996).

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; U33437; AAB40040.1; -. INTERTY). PREVOZESZ; HCV_capsid. InterPro; IPR002522; HCV_capsid. InterPro; IPR002521; HCV_capsid. InterPro; IPR002531; HCV_capsid. Pfam; PF01543; HCV_capsid. Pfam; PF01542; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                    ö
        Score 131; DB 12; Length 191;
Pred. No. 5.5e-12;
4; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87.9%; Score 131; DB 12; Length 191; 82.1%; Pred. No. 5.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tokita H., Okamoto H., Luengrojanakul P., Vareesangthip K., Chainuvati T., Iizuka H., Tsuda F., Miyakawa Y., Mayumi M.; "Hepatitis C virus variants from Thailand classifiable into genotypes in the sixth (6b), seventh (7c,7d) and ninth (9b,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            191
20720 MW; 7A4B84532A1D07F9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1996 (TIEMBLEEL. 01, Created)
01-NOV-1996 (TIEMBLEEL. 01, Last sequence update)
01-NOY-1996 (TIEMBLEEL. 22, Last annotation update)
Genome polyprotein (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              415 AA.
                                                                                                                                                                                                                                               191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4; Mismatches
                                                                                            1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                    5 PKPQRKTKRNTIRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genetic groups.";
J. Gen. Virol. 76:2329-2335(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-Th576;
MEDLINE-96005057; PubMed-7561773;
             87.9%;
82.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity 82.1 tes 23; Conservative
                              Best Local Similarity 82.1
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            081266
                                                                                                                                                                                                                                               968413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
Q81266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                   RESULT 11
Q68413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                   Se propage so se
```

```
4
                                                                                                                                                                                                                                                                                                                                                                                                    Songsivilal S., Kanistanon D., Kunkitti R.; "Identification and characterisation of Thai isolates of hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              virus.";
Submitted (MAR-1995) to the EMBL/GenBank/DDEJ databases.
Submitted the VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GIYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN CAND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ); Score 130; DB 12; Length 43;
b; Pred, No. 1.5e-12;
4; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     43 43 43
43 AA; AB44 MW; Al73B6263AF6DC4F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Last sequence update)
Last annotation update)
                                                   01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
43 AA.
```

us-09-491-146a-32.rspt

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                         SEQUENCE FROM N.A.
STRAIN-RIG292;
Shustov A.V., Gavrilova I.V., Netesov S.V.;
Shustov A.V., Gavrilova I.V., Netesov S.V.;
Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPEID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; AF506622; AAM33398.1; -.
EMBL; AF506622; AAM33398.1; -.
EMBL; PF01543; HCV_capsid,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Other T., Mizokami M., Chstopher T., Ohba K., Suzuki K., Roger W.;
"The nuclectide sequence of the core region of HCV in Pakistan and
Bangladesh: The geographic characterization of HCV in South Asia.";
Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
--- SUBBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL. D29647; BAA21022.1;
PIGHEPRO, 1PRO02522; HCV_capsid.
PÉQUISA3; HCV_capsid. 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87.2%; Score 130; DB 12; Length 109; 85.7%; Pred. No. 4.3e-12; 1ve 2; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                            Score 130; DB 12; Length 73;
Pred. No. 2.8e-12;
4; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
NON_TER 109 109
SEQUENCE 109 AA; 12293 MW; 524E8D425A7E3CB2 CRC64;
                                                                                                                                                                                                                                                                                                                                               76549827A490C896 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    109 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKPQRQTKRNTPRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last seq
01-OCT-2002 (TrEMBLrel. 22, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                73 73
73 AA; 8258 MW;
                                                                                                                                                                                                                                                                                                                                                                            Query Match 87.2%;
Best Local Similarity 82.1%;
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 85.7
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-Banl-1;
                                                                                                                                                                                                                                                                                                  Polyprotein.
                                                                                                                                                                                                                                                                                                             NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              081230
081230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
081230
10 01233
AC 081233
AC 08123
DT 01-NO
DT 01-NO
DT 01-NO
DC COFE
OS HEPAC
OC VITUS
OC HEPAC
OC HEPAC
OC HEPAC
NCB1
RP SEQUE
RC SEGUE
RC SEGUE
RC STRAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
```

Search completed: August 7, 2003, 11:19:03 Job time : 26.6364 secs

```
FRAGMENT TYPE:
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-921-887-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-921-887-32
175, App
175, App
6, Appli
36, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sequence 23, Appl
Sequence 187, App
                                                                                                                                                                                                                               (without alignments)
112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 175,
Sequence 175,
                                                                                                                                                                                                    August 7, 2003, 11:07:41; Search time 10.5455 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
Sequence
Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence Sequence Sequence Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /cgn2_6/ptodata/1/laa/5A_COMB.pep:*
/cgn2_6/ptodata/1/laa/5B_COMB.pep:*
/cgn2_6/ptodata/1/laa/6A_COMB.pep:*
/cgn2_6/ptodata/1/laa/6B_COMB.pep:*
/cgn2_6/ptodata/1/laa/PcTUS_COMB.pep:*
/cgn2_6/ptodata/1/laa/PcTUS_COMB.pep:*
                             GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-921-897-23
US-08-290-665A-188
US-08-290-665A-188
US-08-290-665A-188
US-08-25-10398-188
PCT-US95-10398-188
PCT-US95-10398-190
US-08-836-075A-44
US-08-836-075A-104
US-08-836-075A-104
US-08-836-075A-104
US-08-836-189
US-08-836-189
US-08-836-189
US-08-93-886C-191
US-08-655-886C-191
US-08-655-886C-191
US-08-655-886C-191
US-08-655-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      cal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                        protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         lssued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              seq length: 0
seq length: 2000000000
                                                                                                                                                                                                                                                                                                               US-09-491-146A-32
149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

01000

01000

01000

01000

01000

01000

01000

01000

010000

010000

010000

01000

010000

01000

01000

01000

01000

                                                                                                                                                                                                                                                                                                                                             Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Minimum DB
Maximum DB
                                                                                                                                        OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database :
                                                                                                                                                                                                                                                                                                                                                                          Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Searched:
                                                                                                                                                                                                    Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Result
```

Sequence

```
Sequence 32, Application US/08921887
Sequence 32, Application US/08921887
Patent No. 6030771
GENERAL INFORMATION:
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAM NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
Sequence 3, Appli
Sequence 26, Appli
Sequence 23, Appli
Sequence 13, Appli
Sequence 10, Appli
Sequence 198, Appli
Sequence 14, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 16, Appli
Sequence 2, Appli
Sequence 16, Appli
Sequence 16, Appli
Sequence 17, Appli
Sequence 17, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE: GA
COUNTRY: USA
ZUD: 30303-1769
COMPUTER REDABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0%; Score 149; DB 3;
100.0%; Pred. No. 3.3e-15;
tive 0; Mismatches 0;
US-08-530-550-3

US-08-26-037-26

PCT-US94-07088-9

PCT-US95-13660-3

US-08-836-075A-10

US-08-635-886C-198

US-08-635-886C-232

US-08-635-886C-232

US-08-836-075A-14

US-08-934-977-8

US-08-384-616-8

US-08-904-666A-8

US-08-904-666A-8

US-08-904-666A-8

US-08-904-666A-8

US-08-911-95-2

US-08-810-075A-46

US-08-810-075A-60

US-08-810-075A-60

US-08-810-075A-60
                                                                                                                                                                                                                                                                                                                                                                         ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATTORNEY/ACENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
TELECOMMUNICATION: INFORMATION:
TELEPHONE: 404-818-3700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 maino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          internal
 FILING DATE:
CLASSIFICATION: 435
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE OF INVENTION: NUCLECTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: AMINO ACID SEQUENCES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   84.6%; Score 126; DB 2; Length 191;
llarity 78.6%; Pred. No. 5.8e-11;
Conservative 5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                             Length 28;
                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                      Query Match 84.6%; Score 126; DB 3; L
Best Local Similarity 78.6%; Pred. No. 7.3e-12;
Matches 22; Conservative 5; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
IITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RECISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMONICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 187, Application US/08290665A Patent No. 5882852 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                          LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEX: 421792
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
  23:
                                                                                                                                                                                                                                   ) ORGANISM: Hepatitis virus US-08-921-887-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: MORGAN & FINI
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: homosapiens
INDIVIDUAL ISOLATE: HE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 22; Conserv
                                                                                                                                                                                      ANTI-SENSE: NO ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                      á
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOMARD A.
TITLE OF INVENTION: MOSALC A.
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                          APPLICANT: DELEYS, ROBERT APPLICANT: DELEYS, ROBERT APPLICANT: MARRIENS, Geert APPLICANT: MARRIENS, Geert TITLE OF INVENTION: URBUS PRICE PRICE EPITOPES OF HEPATITIS (TITLE OF INVENTION: UNBER: US/08/635,886C CURRENT APPLICATION NUMBER: US/08/635,886C CURRENT APPLICATION NUMBER: US/08/635,886C PRIOR FILING DATE: 1996-04-25 PRIOR FILING DATE: 1994-10-28 PRIOR FILING DATE: 1994-10-28 PRIOR FILING DATE: 1993-11-04 NUMBER OF SEQ ID MOS: 286 NUMBER OF SEQ ID MOS: 286 SOFTWARE: Patentin version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 137; DB 4; Length 319;
Pred. No. 2.5e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPQRQTKRNTLRRPQNVKFPGGGQIVG 32
                      PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATTORNEY AGENT INFORMATION:
NAME: WARREN WILLIAM L.
REGISTATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                 ; Sequence 211, Application US/08635886C
; Patent No. 6555114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 23, Application US/08921887 Patent No. 6030771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 91.9%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                   GENERAL INFORMATION:
APPLICANT: LEROUX-ROELS, Geert
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     404-818-3799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CITY: Atlanta
STATE: GA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
                                                                                                                                                              US-08-635-886C-211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-635-886C-211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-921-887-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 211
LENGTH: 319
                        ò
                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        임
```

```
Gaps
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS: ADDRESS: ADDRESSE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE CITY: NEW YORK STREET: NEW YORK COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: BUXH, J., MILLER, R.H. AND
PAPLICANT: PURCELL, R.H.

TITLE OF INVENTION: NUCLEORED AND DEDUCED

TITLE OF INVENTION: ORE GENES OF ISOLATES OF HEPATITIS C VIRUS

TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE

TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE

TITLE OF INVENTION: 200UBNCES IN DIAGNOSTIC METHODS AND VACCINES

WHOMBER OF SEQUENCES: 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       84.6%; Score 126; DB 2; 78.6%; Pred. No. 5.8e-11; iive 5; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: PCT/US95/10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 187, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SYSTEM: PC-DOS/MS-DOS WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT N. 473.

ATTORNEY/AGENT N. 473.

REGISTRATION NUMBER: 36,459

REFERENCE/COCKET NUMBER: 2026.

TELECOMMUNICATION INFORMATION:

TELEFRONE: (212) 758-4800

TELERA: (212) 758-4800

TELERA: (212) 751-6849

TELEX: 421792

INFORMATION FOR SEQ ID NO: 190:

SEQUENCE CHARACTERISTICS:

LENGTH: 191 amino acids

TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: NEW YORK
STATE: NEW YORK
COUTRY: USA
2IP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
COMPUTER: OF CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: homosapiens
INDIVIDUAL ISOLATE: DK12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CURRENT APPLICATION DATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY: unknown ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-290-665A-190
                                                                                                                                                                                                                                                                                                                              10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PCT-US95-10398-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HERATITIES C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    84.6%; Score 126; DB 2; Length 191; 78.6%; Pred. No. 5.8e-11; Live 5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUCLEOTIDE AND DEDUCED
                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUC
TITLE OF INVENTION: CORE GENES OF ISOLAT
TITLE OF INVENTION: CORE GENES OF REAGE
TITLE OF INVENTION: SEQUENCES IN DIAGNOS
TITLE OF INVENTION: SEQUENCES IN DIAGNOS
CORRESPONDENCE ADDRESS:
                                     1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-NUG-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MILLER, R.H. AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 190, Application US/08290665A
Fatent No. 5882852
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
TITLE OF INVENTION: NUCLECTIDE AND DE
TITLE OF INVENTION: AMINO ACID SEQUENCY
TITLE OF INVENTION: AMINO ACID SEQUENCY
TITLE OF INVENTION: CORE GENES OF ISO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2026-4116
                                                                                                                                                                                                                                                                                        Sequence 188, Application US/08290665A Patent No. 5882852 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATTORNEY/AGENT INFORMATION:
NAME: RECISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 759-4800
TELETA: 421792
INFORMATION FOR SED NO: 188:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 84.67
Best Local Similarity 78.65
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGINAL SOURCE
                                                                                                                                                                                                                                              US-08-290-665A-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 6
US-08-290-665A-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-290-665A-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STATE: NI
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      , a
                                     ò
                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
```

Gaps

```
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: MAINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: AND ACID SEQUENCES OF HEPATITIS C VIRUS
TITLE OF INVENTION: SEQUENCES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN 6 FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                  84.6%; Score 126; DB 5; Length 191;
1larity 78.6%; Pred. No. 5.8e-11;
Conservative 5; Mismatches 1: Indels
                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. 5.8e
5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION DATA:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION UMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT-US95-10398-190
; Sequence 190, Application PC/TUS9510398
; GENERAL INFORMATION:
                                           TELEPHONE: (212) 758-4800
TELERA: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYRE: amino acid
TYRE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEX: 421792
INFORMATION FOR SEQ ID NO: 190: SEQUENCE CHARACTERISTICS: LENGTH: 191 amino acids TYPE: amino acids STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RECISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 202
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDIUM TYPE: FLOPPY DISK COMPUTER: IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                        INDIVIDUAL ISOLATE: S52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HEW YORK NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: RICHARD W. BORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRANDEDNESS: unknown TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                TOPOLOGY: unknown ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 22; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N.H.

NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCED OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 84.6%; Score 126; DB 5; Length 191; Best Local Similarity 78.6%; Pred. No. 5.8e-11; Matches 22; Conservative 5; Mismatches 1; Indels
  CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTONEY/ARGNT INFORMATION:
NAME: RICHARD W BORK
REGISTRATION NUMBER: 36,459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 188, Application PC/TUS9510398
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: BUKH, J., MILLER, R.H. TITLE OF INVENTION: NUCLEOTIDE AND DEDUCITION OF TITLE OF INVENTION: AMINO ACID SEQUENCES; TITLE OF INVENTION: AND THE USE OF REAGES; TITLE OF INVENTION: AND THE USE OF REAGES; TITLE OF INVENTION: SEQUENCES IN DIAGNOS; NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPKROTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
PPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY: USA
2IP: 10154
COMPUTER READABLE FORM:
COMPUTER: THE FLOPPY DISK
COMPUTER: THE PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                          REFERENCE/DOCKET NUMBER: 2026
RELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELERA: (212) 758-4800
TELERA: 421792
INFORMATION FOR SEQ ID NO: 187: SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE CITY: NEW YORK STAFF: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REGISTRATION NUMBER: 36,459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INDIVIDUAL ISOLATE: HK10 PCT-US95-10398-187
15-AUG-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RICHARD W. BORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PCT-US95-10398-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
APPLICANT: DELEYS, ROBERT
APPLICANT: DELEYS, ROBERT
APPLICANT: DELEYS, ROBERT
APPLICANT: MARRENENS, Geert
TITLE OF INVENTION: INMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VINUS
FILE REFERENCE: 2752-18
CURRENT FILING DATE: 1996-04-25
PRIOR FILING DATE: 1996-04-25
PRIOR FILING DATE: 1994-10-28
PRIOR FILING DATE: 1994-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PATENTIN OFFICIALION NUMBER: BP 93402718.6
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PATENTIN OFFICIALION OFFICIALION 3.1
SEQ ID NO 230
LENGTH: 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 126; DB 4; Length 319;
Pred. No. 1e-10;
5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COCATION: (233]..(233)
CTHER INFORMATION: Xaa is any amino acid
US-08-635-886C-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LOCATION: (144)...(144)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY: MISC_FEATURE
LOCATION: (156)..(157)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ION: (161)..(161)
INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ion: (167)..(167)
INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ION: (171)..(172)
INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ION: (174)..(174)
INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCATION: (177)..(177)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LOCATION: (232)..(232)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any
                                               APPLICANT: LEROUX-ROELS, Geert
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ch 84.6%;
1 Similarity 78.6%;
22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 12
US-08-921-887-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local S
Matches 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: MAERTENS, GEERT
APPLICANT: STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
TITLE OF INVENTION: AGENTS
NUMBER OF SEQUENCES: 207
                                                                                                                                                        0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                           Gaps
                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                        Score 126; DB 5; Length 191;
Pred. No. 5.8e-11;
5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 3; Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READALLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOGTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 126; DB 3;
Pred. No. 1e-10;
5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                  1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                        1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: PCT/EP95/04155
FILING DATE: 23 Oct 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: US/08/836,075A FILING DATE: 21 Apr 1997 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433 CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; Sequence 230, Application US/08635886C
                                                                                                                                                                                                                                                                                                                               US-08-836-075A-44
; Sequence 44, Application US/08836075A
Patent No. 6180768
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       84.6%;
                                                                                                        Query Match
Best Local Similarity 78.6%;
Matches 22; Conservative
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: DK12
PCT-US95-10398-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 84.6
Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: USA
ZIP: 77210-4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-635-886C-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                           g
```

Gaps

```
NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEBATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 125; DB 3; Length 74;
Pred. No. 2.9e-11;
3; Mismatches 2; Indels
                              COMPUTER: TIPPE OF TAILORY OF TAILOR OF TAILOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 189, Application US/08290665A
Patent No. 5802852
GENERAL INFORMATION:
APPLICANT: BUKK, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE: 15-AUG-1994
CLASSFFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: IN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATTORNEY ACCENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 202
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INFORMATION FOR SEC ID NO: 104
SEQUENCE CHARACTERISTICS:
LENGTH: 74 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        83.9%;
82.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: NULLEOT TITLE OF INVENTION: OORE GE TITLE OF INVENTION: OORE GE TITLE OF INVENTION: SEQUENCE: TITLE OF INVENTION: SEQUENCE: CORRESPONDENCE ADDRESS: ADDRESSE: ADDRESS: ADDRESSE: MANAYORK STREET: 345 PARK AVENUE CITY: NEW YORK STATE: NEW YORK COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : (212) 758-4800
(212) 751-6849
                Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 83.9
Best Local Similarity 82.1
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CURRENT APPLICATION DATA:
APPLICATION NUMBER: US,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14
US-08-290-665A-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-836-075A-104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       င်
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 104, Application US/08836075A

Patent No. 6180768

GENERAL INFOMATION:

APPLICANT: MAERTENS, GEERT

APPLICANT: STUYVER, LIEVEN

TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES

TITLE OF INVENTION: AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC

TITLE OF INVENTION: AGENTS

NUMBER OF SEQUENCES: 207

CORRESPONDENCE ADDRESS:

ADDRESSE: ANDLD, WHITE & DURKEE

STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 3; Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY: USA
ZIP: 30303-1769
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMpatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match

83.9%; Score 125; DB 3;
Best Local Similarity 82.1%; Pred. No. 1e-11;
Matches 23; Conservative 3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 24, Application US/08921887
Patent No. 6030771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
CLASSIFICATION: 435
ATORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REPERENCE/COCKET NUMBER: 03065
TELEPHONE: 404-818-370
TELEFAX: 404-818-379
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acids
TYPE: amino acids
STRANDEDNESS: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; ORGANISM: Hepatitis virus US-08-921-887-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N: USA
77210-4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ANTI-SENSE: NO FRAGMENT TYPE: 1 ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                  CITY: Atlanta STATE: GA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-836-075A-104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CITY: H
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
ô
                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N.H.

NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
CORE GENES OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                         .
0
                                                                                                                                                                                                                                          Score 125; DB 2; Length 191;
Pred. No. 8.1e-11;
6; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PUNCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUC
TITLE OF INVENTION: AMINO ACID SEQUENCES
TITLE OF INVENTION: CORE GENES OF ISOLAT
TITLE OF INVENTION: AND THE USE OF REAGE
TITLE OF INVENTION: SEQUENCES IN DIAGNOS
NUMBER OF SEQUENCES: 263
                                                                                                                                                                                                                                                                                                                                                     1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:

NAME: RIGHARD W. BORK

REGISTRATION NUMBER: 36,459
REFERNCE/DOCKET NUMBER: 2026-4116
TELEPHONE: (212) 758-4800

TELEPHONE: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ZUP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 189, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MORGAN & FINNEGAN
TELEX: 421792
INFORMATION FOR SEQ 1D NO: 189:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEX: 421792
INFORMATION FOR SEQ ID NO: 189:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                            Query Match 83.9%;
Best Local Similarity 75.0%;
Matches 21; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 191 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STREET: 345 PARK AVENUE CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INDIVIDUAL ISOLATE: S2
PCT-US95-10398-189
                                                                                                                                                            ORGANISM: homosaplens
INDIVIDUAL ISOLATE: S2
US-08-290-665A-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY: unknown IGINAT. comma
                                                                               TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NEW YORK
USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                          PCT-US95-10398-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE: NE
COUNTRY:
                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                      a
```

Appl

Sequence

Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Seq

```
APPLICANT: MAERTENS, GEERT
STUTVER, LIEVENS SEOURNEES OF HERATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: NEW SEQUENCES OF PROPHYLACTIC, THERAPEUTIC AND DIAGNO
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPTR: 1253

COMPTR: 1253

COMPTR: 1254

ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMFUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/NS-DOS

SOFTWARE: Microsoft Word 6.0 / ASCII text output

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 10S/09/851,138

FILING DATE: 09-May-2001

PRIOR APPLICATION NUMBER: 09/836,075

FILING DATE: 20 OKT 1994

APPLICATION NUMBER: EP 94870166.9

FILING DATE: 20 OKT 1994

APPLICATION NUMBER: EP 95870076.7
US-09-952-572-9
US-09-929-955-1
US-09-747-419-20
US-09-891-894-3
US-10-104-966-1
US-10-259-275-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                              ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                              CORRESPONDENCE ADDRESS:
RANDLD, WHITE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                          Sequence 44, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                            AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                      207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
  809
2894
22894
3011
3011
3011
3011
3011
                                                                                                                                                                          809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                  US-09-851-138-44
   RESULT
   Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Appl1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 104, App
Sequence 1, Appl
Sequence 1, Appl
Sequence 7, Appl
Sequence 77, Appl
Sequence 78, Appl
Sequence 60, Appl
Sequence 152, Appl
Sequence 153, Appl
Sequence 152, Appl
Sequence 153, Appl
Sequence 154, Appl
                                                               August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Description
                                                                                                                                                                                                                                                                                      451899
          GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-851-138-44

US-09-851-138-104

US-09-851-138-104

US-09-758-308-1

US-09-758-308-1

US-09-921-397-77

US-09-921-397-77

US-09-921-397-78

US-09-851-138-14

US-09-851-138-14

US-09-851-138-14

US-09-851-138-14

US-09-891-138-14

US-09-891-138-14

US-09-891-138-14

US-09-899-046-152

US-09-899-046-152
                                                                                                                                                                                         .al number of hits satisfying chosen parameters:
                                                                                                                         1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                        451899 segs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                                Published_Applications_AA:*
                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                protein search, using sw model
                                                                                                                                            BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                            length: 0
length: 2000000000
                                                                                                      US-09-491-146A-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             www.www.ww.ww.ww.
                                                                                                                                                                                                              seq
                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Perfect score:
                                                                                                                                                                                                            Minimum DB
Maximum DB
                                                protein
                                                                                                                                                                                                                                                                                Database :
                                                                                                                          Sequence:
                                                                                                                                                                         Searched:
                                                                  Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Result
                                                                                                       Title
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ટ
                                               δ
```

6 DURKEE

Sequence 23, Appl Sequence 40, Appl Sequence 4, Appl1 Sequence 1, Appl1 Sequence 1, Appl1 Sequence 20, Appl Sequence 3, Appl1 Sequence 1, Appl1 Sequence 20, Appl1 Sequence 20, Appl1 Sequence 20, Appl1

```
STOYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOS
                                                                                                              APPLICANT: DOLANT: MICHEL
APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: LADAVIBER, LAGURENT
APPLICANT: LADAVIBER, LAGURENT
APPLICANT: LADAVIBER, LAGURENT
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTION
FILE REPERBACE: 103959
CURRENT FILING DATE: 1999-09-07.
PRIOR FILING DATE: 1999-09-07.
PRIOR FILING DATE: 1999-09-07.
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PALCHIN VET. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
TITLE: of Hepatitis Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-TOS/MS-DOS
SOFWRARE: MICTOSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match

81.2%; Score 121; DB 15;
Best Local Similarity 78.6%; Pred. No. 3.3e-10;
Matches 22; Conservative 4; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US/09/851,138 FILING DATE: 09-May-2001 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: 08/836,075
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL: Proc. Natl. Acad. Sci. U.S.A. VOLUME: 88 PAGES: 3392-3396 DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 10, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
                      Sequence 1, Application US/10367677
Publication No. US20030118604A1
GENERAL INFORMATION:
APPLICANT: JOLIVET, MICHEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RELEVANT RESIDUES: 2 TO 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: MAERTENS, GEERT
SPELICANT: MAERTENS, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                          Query Match

84.6%; Score 126; DB 10; Length 319;
Best Local Similarity 78.6%; Pred. No. 5.5e-10;
Matches 22; Conservative 5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83.9%; Score 125; DB 10;
82.1%; Pred. No. 1.6e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
RADBRESSES: RNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: US/09/851,138 FILING DATE: 09-May-2001 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-09-851-138-104
                                               TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-851-138-44
                                                                                                                                                                                                                                                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-851-138-104
: Sequence 104, Application US/09851138
:blication No. US20020183508A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                         LENGTH: 319 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 74 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO: 104: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 83.9
Best Local Similarity 82.1
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY: USA
ZIP: 77210-4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

Fri Aug

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    polypeptides selected from a
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81.2%; Score 121; DB 10; Length 97; 78.6%; Pred. No. 7.8e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                            SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 121; DB 10;
Pred. No. 8.3e-10;
4; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GREERAL INCORNATION:
APPLICANT: HYBRIGENICS
TITLE OF INVENTION: SID nucleic acids and poly
TITLE OF INVENTION: pathogenic strain of the
TITLE OF INVENTION: applications thereof
FILE REFERENCE: B4809A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
LENGTH: 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                            CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: 0S 08/259,721
FILING DATE: 29-AUG-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
  MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 14, Application US/09851138; Publication No. US20020183508A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 77, Application US/09921397
Patent No. US20020151484A1
                                                                                                                                                                                                                                                                                                                                                                   TELECOMMUNICATION INFORMATION:
TELEPHONE: (2027) 793-6040
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                               REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 78.69
----hes 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 97 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               amino acid
                                                                                                                                    FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 1, Application US/09758308
Patent No. US20020090607A1
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
FILE OF INVENTION: PROTEINS
FILE REFERENCE: 14114.034902
CURRENT APPLICATION NUMBER: US/09/758,308
CURRENT FILING DATE: 2001-01-10
                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
0
                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match 81.2%; Score 121; DB 10; Length 74; Best Local Similarity 78.6%; Pred. No. 5.8e-10; latches 22; Conservative 4; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: PIKE, IAN
TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 701-E Columbia Square
STREET: 555 13th Street, N. W.
CITY: Washington
STATE: D. C.
COUNTRY: U. S.
ZIP: 20004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nery Match 81.2%; Score 121; DB 9; Jest Local Similarity 78.6%; Pred. No. 7.3e-10; Matches 22; Conservative 4; Mismatches 2
FILING DATE: CURNOWN>
PULLIATION NUMBER: EP 94870166.9
FILING DATE: 21 oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                           NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                               TOPOLOGY: linear;
HOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MBER: 60/092,339
1999-07-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 8, Application US/09756875 Patent No. US20020150990A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                  LENGTH: 74 amino acids TYPE: amino acid
                                                                                                         ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PatentIn version 3.0
                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; ORGANISM: Hepatitis C Virus US-09-758-308-1
                                                                                                                                                                                           INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION NUMBER:
PRIOR FILING DATE: 1999-(
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQ ID NOS:
SOFTWARE: Patentin v.
SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 5
US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
```

```
RESULT 10
US-09-851-138-46

Sequence 46, Application US/09851138

Publication No. US020183508A1

GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT

TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES

TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES

AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STUTUE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOS
                            Gaps
                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOFTWARE: Microsoft Word 6.0 / ASCII text output CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                     2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 121; DB 10;
Pred. No. 1.1e-09;
4; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                     Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                   111:1:1:1111 111::1111 111111 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                             1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPKROTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STATE: TEXAS.
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 60, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 137 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHI
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE: 28 Jun 19
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER: IBM PC C
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: linear
                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity
Matches 22; Conserva
                 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-851-138-46
                     Matches
                                                                                                g
                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               polypeptides selected from the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 113;
                                                                                                                                                                                                                                                                                                    COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
A APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 121; DB 10;
Pred. No. 8.8e-10;
4; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 121; DB 10;
Pred. No. 9.2e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: HYBRIGENICS
TITLE OF INVENTION: SID nucleic acids and pol-
TITLE OF INVENTION: Bathogenic strain of the
TITLE OF INVENTION: pathogenic strain of the
TITLE OF INVENTION: applications thereof
FILE REFERENCE: BASONA - 3AZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
                                                                                          NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
SIREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
PAPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                            STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 78, Application US/09921397
Patent No. US20020151484A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 108 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 78.6%;
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQ ID NO 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

Gaps

ö

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PAtentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81.2%; Score 121; DB 11; Length 166; 78.6%; Pred. No. 1.4e-09;
                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/899,046
Pred. No. 1.4e-09;
4; Mismatches 2;
                                                                  1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                              1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; Sequence 152, Application US/09878281
; Publication No. US20030032005A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; Sequence 42, Application US/09899046; Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
Similarity 78.6%;
22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
INFORMATION FOR SEQ ID NO: 4
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 22; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                     US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY:
  Best Local
Matches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
TITLE OF INVENTION: 270
COMPUTER READABLE FORM:
MEDIOM TYPE: RIOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81.2%; Score 121; DB 10; Length 138; 78.6%; Pred. No. 1.1e-09;
                                                                                                                                ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 11;
                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: 08/836,075
FILING DATE: 
ATORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4; Mismatches
                       WHITE & DURKEE
                                                                                                                                                                                                                                                                    CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81.2%; Score 121;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |||:|:||||| ||||| PKPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Sequence 152, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 138 amino acids
                         ADDRESSEE: ARNOLD, WH
STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO: 152: SEQUENCE CHARACTERISTICS:
    CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protein
                                                                  CITY: HOUSTON STATE: TEXAS
                                                                                                                COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
```

Gaps

us-09-491-146a-32.rapb

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                             TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
APPLICATION NUMBER: US/09/899,046
                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .;
0
                                              ö
Query Match 81.2%; Score 121; DB 11; Length 169; Best Local Similarity 78.6%; Pred. No. 1.4e-09; Matches 22; Conservative 4; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 121; DB 11; Length 169;
Pred. No. 1.4e-09;
4; Mismatches 2; Indels
                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                            5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                  NESULT 15
US-09-899-046-44
Sequence 44, Application US/0989046;
Publication No. US20030008274A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 169 amino acids

TYPE: amino acid

TOPOLOGY: linear

WOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 81.2%;
Best Local Similarity 78.6%;
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                   ò
                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

Search completed: August 7, 2003, 12:01:14 Job time: 14.3636 secs

5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32

q

Scoring table:

Searched:

2000

Minimum I Maximum I

Database

score:

Title: Perfect :

Sequence:

protein

Run on:

New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual

Khudyakov YE;

Fields HA,

WPI; 1999-204671/17.

HCV capsid peptide Non-A, non-B hepat Hepatitis C virus Hepatitis C virus

AAR30689 AAR54065 AAW37380 AAW66083

AAR30688

100.0 92.7 92.0 92.0 92.0 92.0

139 138 138 138 138 138

Length

Ouery Match 1

Score

Result Š CN14 fragment of H HCV capsid peptide HCV capsid peptide

Fri Aug

X S

```
The invention relates to a mosaic protein, comprising a plurality of annologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences AAVO6673-681 represent amino acid sequence of each monomer
                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA and amino acid sequence of HCV envelope 1 and core proteins -used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                    28;
                                                                                                                                                                                                                                                                                                                                                    Length
                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                100.0%; Score 150; DB 20;
100.0%; Pred. No. 2.2e-14;
ilve 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus isolate SA6 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                         (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 4; Page 221-222; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR92985 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                             comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Purcell RH;
                  Claim 5; Fig 9; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                Local Similarity 100.
les 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1996-139709/14.
N-PSDB; AAT16659.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                               28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO9605315-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22-FEB-1996,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    hepatitis.
                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                  uery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  E1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR92985
                                                                                                                                                                                                                                                                                                                                                                    ast Loc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 2
AAR92985
                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This sequence is encoded by the oligonucleotide, CN14, and represents the peptide fragment CP14. CP14 is a fragment of the core region of hepatitis C virus (HCV). CP14 may be used in the detection of HCV infection and to raise antibodies against it.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Detection of hepatitis C virus - using oligopeptide fragment of HCV
                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus; HCV; open reading frame; "common" seguence;
                         Length 191;
                                                                                                                                                                                                                                           CN14; CP14; core region; hepatitis C virus; HCV; detection;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 36;
                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                      Score 139; DB 17;
Pred. No. 6.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB 16;
Pred. No. 1.5e-12;
2; Mismatches 1;
                                             2; Mismatches
                                                                      28
                                                                                5 PKPQRKTQRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                                                    Ą.
                                                                                                                                                                                                                       CN14 fragment of HCV core region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5
                                                                                                                                                  AAW06487 standard; Protein; 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 1; Page 6; 7pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR30687 standard; peptide; 38
                      92.7%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                         93JP-0156026.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV capsid peptide No. 23.
                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                               (IMMO ) IMMUNO JAPAN KK.
                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1995-047903/07.
N-PSDB; AAT45055.
                             Local Similarity
es 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 25; Conserv
191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        36 AA;
                                                                                                                                                                                                                                                                                                    JP06327482-A.
                                                                                                                                                                                                                                                                                                                                                  21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                        21-MAY-1993;
                                                                                                                                                                                                31-JAN-1997
                                                                                                                                                                                                                                                                                                                           29-NOV-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     core region
                                                                                                                                                                                                                                                                              Synthetic.
 Sequence
                                                                                                                                                                         AAW06487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR30687;
                                                                                                                                                                                                                                                         antibody.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                      Query Match
                                             Matches
                                                                                                                            RESULT 3
                                                                                                                                      AAW06487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                         q
                                                                      à
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
```

capsid non-A,

```
Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope groups. These peptides can be used in immunoassays for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
                                                                                                                        Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                               sequences given in AAR30665-89 represent fragments of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.0%; Score 138; DB 14;
89.3%; Pred. No. 1.6e-12;
tive 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                              (Updated on 25-MAR-2003 to correct PN fleld.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Disclosure; Fig 1F; 66pp; English
                                                   Todd JA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Todd JA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR30689 standard; peptide; 38
                  (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92WO-US03635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91US-0714471
91US-0718052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 89.39
...+Ahas 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV capsid peptide No. 25.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-MAR-2003 (updated)
11-MAY-1993 (first eni
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Leahy DC,
                                                   Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1993-018073/02.
                                                                                      WPI; 1993-018073/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                 38 AA;
                                                                                                                                            virus - for immu
non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO9222571-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23-DEC-1992.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Jolley ME,
                                                     Jolley ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR30689;
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                           The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope groups. These peptides can be used in immunoasasys for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .
0
                                                                                                                                                                                                                                                                                                                                                       Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB 14; Length 38;
Pred. No. 1.6e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                (BAXT ) BAXTER DIAGNOSTICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR30688 standard; peptide; 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                          91US-0714471.
91US-0718052.
                                                                                                                                                           92WO-US03635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91US-0714471
91US-0718052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92WO-US03635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV capsid peptide No. 24.
                                                                                                                                                                                                                                                                                    Jolley ME, Leahy DC,
                                                                                                                                                                                                                                                                                                                    WPI; 1993-018073/02
                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  W09222571-A1.
                                                                                      W09222571-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-JUN-1991;
20-JUN-1991;
                                                                                                                                                         29-APR-1992;
                                                                                                                                                                                                            20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        29-APR-1992;
                                                                                                                                                                                              13-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       23-DEC-1992
```

Sequence

TEXEXOCOCOCOCOXX

AAR30688;

3ULT 5 .R30688

ô g

ö

Gaps

ő

```
The present sequence represents a Hepatitis C virus (HCV) protein sequence from the disclosure of the present specification. The present specification describes a chimeric HCV peptide antigen which comprises at least 2 peptide epitope regions from the HCV polypeptide core region, 2 peptide epitope regions from the NS3 region and at least 2 peptide epitope regions from the NS3 region and at specifically with an antibody produced by a human infected by HCV. The peptide can detect a wide range of HCV infections with high sensitivity.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    New chimeric peptide antigen derived from hepatitis C virus protein - useful for detecting HCV infections
                                                                                                                                                                  Hepatitis C virus; HCV; chimeric; antigen; detection; core region;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        antigenic; nucleocapsid; p21 protein; alpha-helical; immunogen; antibody; hepatitis C virus; epitope; N-terminal fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus p21 protein N-terminal fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.0%; Score 138; DB 19;
89.3%; Pred. No. 1.8e-12;
11ve 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTERNTNRRPQDVRFSGGGGIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Page 24; 30pp; Japanese.
                                                                                                                                 Hepatitis C virus C-1 protein 1-43.
                         AAW37380 standard; Protein; 43 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAW66083 standard; peptide; 44
                                                                                                                                                                                    infection
                                                                                                                                                                                                                                                                                                                             97JP-0027015.
                                                                                                                                                                                                                                                                                                                                                              96JP-0024045.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 89.3%,
And 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98WO-FR00442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97FR-0002878
                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16-NOV-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1998-022248/03
                                                                                                                                                                                    epitope; NS3; NS4;
                                                                                                                                                                                                                                                                                                                                                                                                 (TOFU ) TONEN CORP.
                                                                                                                                                                                                                       Hepatitis c virus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hepatitis c virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  43 AA;
                                                                                                                                                                                                                                                                                                                          10-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                            09-FEB-1996;
                                                                                                                                                                                                                                                      JP09278794-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   05-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     05-MAR-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO9839360-A1
                                                                                                11-MAR-1998
                                                                                                                                                                                                                                                                                          28-OCT-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11-SEP-1998
                                                              AAW37380;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW66083;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                        Gaps
The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "cormon" sequence. These peptides are contained in the capsid protein of the virus and themselves contained in the capsid protein of the virus and themselves contain epitope antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DNA coding a Non-A, non-B hepatitis virus antigen - useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus #4 and #6 genes were isolated (AAQ64068-69). Both genes contain the core, ENV, NSI, NS2 and NS3 regions. A core region fragment is given in AAQ64067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Non-A, non-B hepatitis virus; NANBHV; hepatitis C virus; HCV; core; ENV; NS1; NS2; NS3; antigen; detection.
                                                                                                                                                                                                                                  Score 138; DB 14; Length 38;
Pred. No. 1.6e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Non-A, non-B hepatitis virus coreI region fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     92.0%; Score 138; DB 15;
89.3%; Pred. No. 1.6e-12;
ive 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (TOKR-) 2H TOKYOTO RINSHO IGAKU SOGO KENKYUSHO
                                                                                                                                                                (Updated on 25-MAR-2003 to correct PN fleld.)
                                                                                                                                                                                                                                                                                                       1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                           5. PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPORKTERNTNRRPODVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; Page 11; 22pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (SANW ) SANWA KAGAKU KENKYUSHO CO. (TOFU ) TONEN CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                              AAR54065 standard; Protein; 38
                                                                                                                                                                                                                         Query Match
Best Local Similarity 89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92JP-0088140.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92JP-0088140.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           detecting HCV within serum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 89.3 Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1994-205028/25.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  38 AA;
                                                                                                                                                                                                   38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAQ64067
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JP06141870-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14-FEB-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-MAY-1994.
                                                                                                                                                antibodies
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR54065;
                                                                                                                                                                                                   Sequence
```

AARS4065

요

ò

Query Match

ò

XTTXXXXXXXXXXXXXX

RESULT 8

ö

Gaps

ö

Length 43 Indels S

```
The invention relates to a peptide, which is recognised by antibodies against amino acids 2.45 at the N-terminus of the core (or nucleocapsid) p21 protein of hepatitis C virus (HCV), or its variants. The peptide has a tertiary structure consisting of two alpha-helical fragments, almost perpendicular to each other in space, connected by a junction peptide. Excluded are all proteins and peptides comprising, or consisting of, the N-terminal part of p21 (starting from amino acid 1 or 2). Also new are (1) monoclonal or polyclonal antibodies produced using the peptide as an immunogen and (2) complex consisting of the peptide specifically bound to some other molecule, particularly peptide or nucleotide fragments or functionalised aromatic compounds. The peptide is used (1) as immunogen for generating and unithodies and (1) for detecting and quantifying either antibodies against p21 or HCV-derived mRNA (by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              prevent HCV infections. The present sequence represents the N-terminus the core p21 protein of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Monoclonal antibodies useful for detecting and/or quantifying hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                 quantifying either antibodies against p21 or HCV-derived mRNA (by complex formation). Antibodies are used correspondingly to detect HCV crelected antigens. The peptide and antibodies may also be used to treat.
                                                                                                                                       nseq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                  New peptide from the N-terminus of hepatitis C virus p21 protein containing the immunodominant epitope - and related antibodies, tfor diagnosis, treatment and prevention of hepatitis C infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               hepatitis C virus; core protein; monoclonal antibody;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                   Penin F;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Jolivet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 138; DB 19;
Pred. No. 1.8e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus Core protein amino acids 2-45.
                                                 Ladaviere L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Paranhos BC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               diagnosis; infection; sandwich immunoassay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTERNINRRPQDVRFSGGGQIVG
                                                                                                                                                                                     Disclosure; Page 16; 37pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                 Lacoux X,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY26952 standard; peptide; 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Yvon S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98FR-0003087.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98FR-0003087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 92.0 Best Local Similarity 89.3 Matches 25; Conservative
                                                 Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Piga N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              virus core protein
                (INMR ) BIO MERIEUX
                                                                                WPI; 1998-495793/42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1999-530397/45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   44 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FR2775690-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Jolivet RC,
                                                Dalbon P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY26952;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Epitope;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 1
                                                                                                                                                                                                                                                                                                                                                                  `888888888%&
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           XXX
```

ö

or

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The present peptide, designated $426, corresponds to residues 2 to 45 of the N-terminus of the human hepatitis C virus (HCV) core protein. It is an immunodominant region containing conformational type epitopes and linear type epitopes. It manifests an immunoraactivity with all the sera of individuals or blood samples infected with HCV and which possess antibodies directed against the core protein. An amino acid may be substituted for homologous amino acids and side chains and peptide bonds may be modified. For example, L-amino acids may be replaced by D-amino acids, amine groups may be acetytated, and so on. The native antigenic sequence and its antigenic derivatives may be used for detection of hepatitis C virus and for raising antibodies against the virus.
                                                               nsed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polypeptides that bind to anti-hepatitis C virus antibodies, useful for diagnosing and preventing hepatitis C infections
                                         Peptides AAY26949-Y26955 represent peptide epitopes derived from the N-terminus of the hepatitis C virus core protein. The peptides are used to generate monoclonal antibodies or antibody fragments specific for hepatitis C virus (HCV) core protein. The monoclonal antibodies are used for early diagnosis of HCV infections, especially by two-antibody
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Human; hepatitis C virus; HCV; immunogen; antigen; antibody; virucide; hepatotropic; anti-inflammatory; virus detection; vaccine.
                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Human hepatitis C virus core protein N-terminus, residues 2-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                          Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 44;
                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ပဲ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Jolivet - Reynaud
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 138; DB 21;
Pred. No. 1.8e-12;
2; Mismatches 1;
                                                                                                                                                                                          20;
                                                                                                                                                                                          Score 138; DB 20;
Pred. No. 1.8e-12;
2; Mismatches 1
                                                                                                                                                                                                                                                        PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                       Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                        Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 1; Page 42; 50pp; English.
               Claim 3; Page 10; 19pp; French.
                                                                                                                                                                                                                                                                                                                                                                     AAY94410 standard; peptide; 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.0%;
89.3%;
                                                                                                                                                                                        Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99WO-IB01933,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98US-0196155
                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ρ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Dalbon
                                                                                                                            sandwich immunoassay
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-411934/35.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus
                                                                                                                                                          44 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            44 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200031130-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                   11-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  02-JUN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Dalbon P,
                                                                                                                                                                                                                                                                                                                                                                                                     AAY94410;
                                                                                                                                                            Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                       RESULT 11
                                                                                                                                                                                                                                                                                                                                                      AAY94410
X & X & C & C & C & C & X & S
                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                    a
```

C; Virus; HCV; antibody;

Takahashi K;

g

õ

```
The sequences given in AAR25929-30 and AAR29534-42 represent various clones of the polypeptide core-envelope region of the Hepatitis C Virus (HCV) gene of the invention. These sequences were isolated from the serum of a patient suffering from hepatitis C (HC). The RNA sequences encoding these peptides were converted into cDNA using transcriptase in the presence of one of the primer sequences given in AAQ3540-46. The cDNA sequences isolated represent different alleles of the same region of the HCV gene. The entire HCV protein (see AAR29527) is useful in the development of a diagnostic method which is more accurate and effective than conventional ones, in the detection of antibodies raised against a wide range of HCVs which have been hardly detected before. The complete gene may be used in convening system for a substance capable of specifically supressing or controlling a proteolytic processing of a precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Clone; polypeptide; core-envelope; region; Hepatitis C; Virus; HCV; HC; cDNA; transcriptase; primer; allele; diagnostic; antibody;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       New hepatitis C virus gene and its encoded protein – used for diagnosing and vaccinating against hepatitis C virus infections
                 Clone, polypeptide, core-envelope, region, Hepatitis HC; cDNA, transcriptase, primer; allele; diagnostic, supress; control; proteolytic processing; precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.0%; Score 138; DB 13;
89.3%; Pred. No. 2.3e-12;
iive 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polypeptide of HCV. (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                Seki M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Honda Y, Murakami T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ą.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 5; Page 74; 305pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR29535 standard; Protein; 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV core-envelope peptide N1-2.
                                                                                                                                                                                                                                                                                                                                                       (MITU ) MITSUBISHI KASEI CORP
                                                                                                                                                                                                          92EP-0109812.
                                                                                                                                                                                                                                                                              91JP-0287008
91JP-0332329
                                                                                                                                                                                                                                                                                                                     92JP-0099957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 92.0
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1992-417213/51.
                                                                                           Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAQ32444.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      55 AA;
                                                                                                                                                                                                        11-JUN-1992;
                                                                                                                                                                                                                                                                                                                                                                                                               Teranishi Y;
                                                                                                                                                                                                                                                                              07-0CT-1991;
16-DEC-1991;
                                                                                                                                                                                                                                                                                                                     20-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25-MAR-2003
26-APR-1993
                                                                                                                                EP518313-A2
                                                                                                                                                                     16-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                              Hayashi N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR29535;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Clone;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR29535
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The present sequence is the 45 N-terminal amino acids of the core protein of the human hepatitis C virus (HCV). The sequence is an immunodominant region containing conformational type epitopes and linear type epitopes. It manifests an immunoreactivity with all the sera of individuals or blood samples infected with HCV and which possess antibodies directed against the core protein. An amino acid of the present sequence may be substituted for homologous amino acids and side chains and peptide bonds may be modified. For example, L-amino acids may be replaced by D-amino acids, amine groups may be acetylated, and so on. The present antigenic sequence and its antigenic derivatives may be used for detection of hepatitis C virus and for raising antibodies against
                                                                                                                                                                                                                                                                         Human; hepatitis C virus; HCV; immunogen; antigen; antibody; virucide; hepatotropic; anti-inflammatory; virus detection; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polypeptides that bind to anti-hepatitis C virus antibodies, useful diagnosing and preventing hepatitis C infections .
                                                                                                                                                                                                                                       Human hepatitis C virus core protein N-terminus, residues 1-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92.0%; Score 138; DB 21; 89.3%; Pred. No. 1.9e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Jolivet-Reynaud C;
28
                   1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PKPQRKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 2; Page 42; 50pp; English.
                                                                                                                          AAY94409 standard; peptide; 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV core-envelope peptide N1-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR29534 standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                99WO-IB01933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0196155
                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2000-411934/35.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         45 AA;
                                                                                                                                                                                                                                                                                                                                                                      WO200031130-A1
                                                                                                                                                                                                    11-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                              19-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-MAR-2003
26-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                           02-JUN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dalbon P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   for detect
the virus.
                                                                                                                                                             AAY94409;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR29534;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 13
                                                                                                       AAY94409
IID AAY9
XX AC AAY9
XX AL AAY9
DJT 11-S
XX XX
XX XX
XX XX
XX XX
XX Huma
XX Huma
XX XX Huma
XX XX Huma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR29534
```

ö

Gaps

ö

Length 55; Indels

200

ð

EX DAX SX E

supress; control; proteolytic processing; precursor

```
1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Search completed: August 7, 2003, 11:14:09
Job time : 38.6364 secs
                 EP518313-A2.
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The sequences given in AAR25929-30 and AAR29534-42 represent various clones of the polypeptide core-envelope region of the Hepatitis C clones of the polypeptide core-envelope region of the Hepatitis C virus (HCV) gene of the invention. These sequences were isolated from the serum of a patient suffering from hepatitis C (HC). The RNA sequences encoding these peptides were converted into cDNA using transcriptase in the presence of one of the primer sequences given in AAG32540-46. The CNA sequences isolated represent different alleles of the same region of the HCV gene. The entire HCV protein (see AAR29527) is useful in the development of a diagnostic method which is more accurate and effective than conventional ones, in the C detection of antibodies raised against a wide range of HCVs which have been hardly detected before. The complete gene may be used in an in vitro screening system for a substance capable of specifically substance of the processing of a precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Clone; polypeptide; core-envelope; region; Hepatitis C; Virus; HCV; HC; cDNA; transcriptase; primer; allele; diagnostic; antibody; supress; control; proteolytic processing; precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                      New hepatitis C virus gene and its encoded protein - used for diagnosing and vaccinating against hepatitis C virus infections
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.0%; Score 138; DB 13; Length 55; 89.3%; Pred. No. 2.3e-12; Live 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                            Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polypeptide of HCV. (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPORKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 33; Page 87; 305pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV core-envelope peptide S1-3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR29536 standard; Protein; 55
                                                                                                                                                                                                                                                                                        (MITU ) MITSUBISHI KASEI CORP
                                                                                                                                                                                               91JP-0172794.
91JP-0287008.
                                                                                                                                                                                                                                91JP-0332329.
                                                                                                                                          92EP-0109812
                                                                                                                                                                              91JP-0139268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            yz.v¤
Jest Local Similarity 89.39
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (updated)
                                                                                                                                                                                                                                                                                                                                                                                WPI; 1992-417213/51.
N-PSDB; AAQ32445.
                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   55 AA;
                                                                                                                                                                                                                                                                                                                                               Teranishi Y;
                                                                                                                                          11-JUN-1992;
                                                                                                                                                                                                                                  16-DEC-1991;
20-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25-MAR-2003
26-APR-1993
                                                                                                                                                                              11-JUN-1991;
                                                                   EP518313-A2
                                                                                                     16-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR29536;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 15
AAR29536
```

ò g

```
The sequences given in AAR25929-30 and AAR29534-42 represent various clones of the polypeptide core-envelope region of the Hepatitis C clones of the polypeptide core-envelope region of the Hepatitis C virus (HCV) gene of the invention. These sequences were isolated from the serum of a patient suffering from hepatitis C (HC). The RNA cequences encoding these peptides were converted into CDNA using transcriptase in the presence of one of the primer sequences given in AAO32540-46. The CDNA sequences isolated represent different in AAO32507. Is useful in the development of a diagnostic method which is more accurate and effective than conventional ones, in the detection of antibodies raised against a wide range of HCVs which catection of antibodies raised against a wide range of HCVs which can be hardly detected before. The complete gene may be used in an in vitro screening system for a substance capable of specifically supressing or controlling a proteolytic processing of a precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New hepatitis C virus gene and its encoded protein – used for diagnosing and vaccinating against hepatitis C virus infections
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                           Honda Y, Murakami T, Seki M, Takahashi K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.0%; Score 138; DB 13;
89.3%; Pred. No. 2.3e-12;
vative 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polypeptide of HCV. (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 33; Page 88-89; 305pp; English.
                                                                                                                                                                                                                                                                                                                       (MITU ) MITSUBISHI KASEI CORP
                                                                                                                                                           91JP-0172794.
91JP-0287008.
91JP-0332329.
                                                                92EP-0109812.
                                                                                                                             91JP-0139268
                                                                                                                                                                                                                                                              92JP-0099957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1992-417213/51.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAQ32446
                                                                                                                                                                                                                                                                                                                                                                                        Hayashi N, |
Teranishi Y;
                                                             11-JUN-1992;
16-DEC-1992.
                                                                                                                             11-JUN-1991;
                                                                                                                                                                                          07-OCT-1991
16-DEC-1991
                                                                                                                                                                                                                                                              20-APR-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
```

polyprotein polyprotein

Sequence:

Run on:

Searched:

Minimum I Maximum I

```
Secies: hepatitis C virus

N; Contains: core protein
C; Species: hepatitis C virus
A; Aariety, genotype 2, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 841353
R; Van Doorn, L.J; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Accession: 841353
A; Molecule type: genomic RNA
A; Residues: 1-108 </NA>
A; Residues: 1-108 </NA>
A; Residues: C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-108/Product: core protein #status predicted 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nicontains: core protein
C;Species: hepatitis C virus
A;Varitet; genotype 2, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41355
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41355
A;Accession: S41355
                                                                                                                                                  genome
genome
genome
genome
                                                                                                                                                                                                                                                                                                     genome
                                                                                    genome
genome
genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                                                                                                                                                                                                                                                           genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C virus (genotype 2, N2) (fragment)
                                              genome
                                                                genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 138; DB 2; Length 108;
Pred. No. 2.1e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 2; Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Experimental source: genotype 2, N4 C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F; 1-108/Product: core protein *status predicted <NAT>
                                                                                                                                                                                                                                                                                                                                                                                         ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 138;
                                          GNWVCJ
S18030
SAWYVCH
SGNWVCD
JC5620
JO1303
SGNWVJ8
S41351
S41352
S41354
JO1584
S21336
S41365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Molecule type: genomic RNA
A;Residues: 1-108 <VAN>
A;Cross-references: EMBL:229458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
876
3010
3010
3010
3011
3011
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
Matches 25; Conserv
992.0
992.0
992.0
992.0
992.0
992.0
992.0
992.0
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
993.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
903.3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein
polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein
polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     hepa
hepa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein
genome polyprotein
                                                                                                                                            August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polyprotein -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qenome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Jenome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          депоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qenome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome
                 GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                         al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                             PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                     283308 segs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                        protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S12707
PC1284
JQ1925
JQ1926
                                                                                                                                                                                                                                                                                                                                           Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          length: 0
length: 2000000000
                                                                                                                                                                                                                                  US-09-491-146A-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR_76:*
!: pir1:*
?: pir2:*
3: pir3:*
!: pir4:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                seq
                                                                                                                                                                                                                                                      Perfect score:
                                                                                                                                                                                                                                                                                                                         Scoring table:
                                                                                                      OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08
08
08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Database :
```

Result

Š

ö

Gaps

```
Accession: S41341
                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S41348
genome polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
N;Confrains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: 341348
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
Persoription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
Ference number: $41341
ccession: $41348
                                                                                                                                                                                                                                 N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 2, N6
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                        5 by LiPA
                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
N;Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92.0%; Score 138; DB 2; Length 108; 89.3%; Pred. No. 2.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1; Indels
                        Indels
                                                                                                                                                                                                                                                                                                                                                                                          Ç
                                                                                                                                                                                                                                                                                                                    C; Accession: $41357
R; van Doorn, L,J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1
A: Reference number: $41341
Pred. No. 2.1e-12;
2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ... residues: 1-108 cVaN>
A;Cross-references: BMBL:229460
A;Experimental source: genotype 2, N6
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A: Experimental source: genotype 1, N6
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 138; DB 2;
Pred. No. 2.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. 2.16
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2; Mismatches
                                                                 1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                     1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPORKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.0%;
ilarity 89.3%;
Conservative
89.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
A; Cross-references: EMBL: 229451
  Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity 89.3
ses 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S41371
genome polyprotein - hepatit:
N;Contains: core protein
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                    olecule type: genomic RNA residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           윱
                                                                 ö
                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
```

```
A; Variety: genotype 5, N5
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 841371
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
S; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
A; Pescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: 841371
A; Reference number: 841371
A; Residues: 1-112 < VANN
A; Residues: 1-112 < VANN
A; Residues: 1-112 < VANN
A; Residues: 1-112 < Consering EMBL: 229474
A; Experimental source: genotype 5, N5
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-112/Product: core protein #status predicted < MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N.Contains: core protein
C.Species: hepatitis C virus
A:Variety: genotype 1, VI
C:Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Strain Corn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A.Reference number: S41341
A.Rocession: Cype: genomic RNA
A.Residues: 1-112 <VAN>
A.Cross-references: EMBL: Z29444; NID: g443850; PIDN: CAA82582.1; PID: g443851
A.Experimental source: genotypel, NI
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Keywords: capsid protein; core protein: polyprotein
F;1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: genomic RNA
A; Molecule type: genomic RNA
A; Residues: 1-114 <VAN>
A; Cross-references: EMBL: 229473; NID:g443908; PIDN:CAA82611.1; PID:g443909
A; Experimental source: genotype 5, N4
C; Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Lipa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein - hepatitis C virus (genotype 1, N1) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Submitted to the EMBL Data Library, Brouwer, J.T. submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to A; Reference number: $41341
A; Molecuit.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 138; DB 2;
Pred. No. 2.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92.0%; Score 138; DB 2; 89.3%; Pred. No. 2.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTKRNINRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 89.3
Matches 25; Conservative
```

Fri

ò g

```
Local Similarity
hes 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C; Accession: 541344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: S41344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 12
S41350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (genotype 5, N3) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N3
C;Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Rivan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41341
A; Reference number: S41348
A; Rolecule: type: genomic RNA
A; Residues: 1-114 </P>
A; Residues: 1-114 </P>
A; Reperimental source: genotype 5, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Reywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A.Reference number: $41341
A.Accession: $41369
A.Molecule type: genomic RNA
A.Residues: 1-114 <VAN>
A.Cross-references: EMBL: Z29472; NID: g443906; PIDN: CAA82610.1; PID: g443907
A.Experimental source: genotype 5, N3
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Keywords: capsid protein; core protein; polyprotein
F.1-114/Product: core protein #status predicted <MAT>
                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Accession: $41369
an Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Lomitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein - hepatitis C virus (genotype 5, N2) (fragment)
                                                                             Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 2; Length in Pred. No. 2.2e-12;
                                                                                                                            1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 138; DB 2; L4
Pred. No. 2.2e-12;
2; Mismatches 1;
C;Keywords: capsid protein; core protein; polyprotein F;l-114/Product: core protein #status predicted <MAT>
                                                                          Score 138; DB 2;
Pred. No. 2.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Vicontains: core protein

C;Species: hepatitis C virus

C;Species: penotype 5, N2

variety: genotype 5, N2

)ate: 19-May-1994 #sequence_revision 26-Jul-1996
                                                                                                                            2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
                                                                                                                                                                                                          5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                               1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                          Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Accession: $41368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 9
```

ò g

```
N.Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N10
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41342
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, NB
C; Date: 19-May-1994 #sequence_revision 26-Jul-1995 #text_change 17-Nov-2000
C; Accession: S41350
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N.Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N2
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                      A;Accession: S41342
A;Molecule type: genomic RNA
A;Residues: 1-115 <VAN>
A;Cross-references: EMBL:229445; NID:g443852; PIDN:CAA82583.1; PID:g443853
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Cross-references: EMBL:Z29447; NID:g443856; PIDN:CAA82585.1; PID:g443857
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
genome polyprotein - hepatitis C virus (genotype 1, N10) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome polyprotein - hepatitis C'virus (genotype 1, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (genotype 1, N8) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             рy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  92.0%; Score 138; DB 2; Length 115; 89.3%; Pred. No. 2.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ស
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1: Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genorypes 1 to
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                          A:Experimental source: genotype 1, N10
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein #status predicted 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Experimental source: genotype 1, N2 (5.5) Superfamily: hepatitis C vitus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F;1-115/Product: core protein *status predicted 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          92.0%; Score 138; DB 2;
89.3%; Pred. No. 2.2e-12;
iive 2; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. 2.26
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPORKTKRNINRRPODVKFPGGGOIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPORKTERNTNRRPODVRFSGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPORKTERNINRRPQDVRFSGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match
Best Local Similarity 89.37
Best Local Similarity 25. Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: genomic RNA A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: S41350
A; Molecule type: genomic RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Reference number: S41341
```

ö

ö

RESULT 10

δ g

```
A;Variety: genotype 1, N5
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
A; Residues: 1-115 <VAN>
A; Residues: 1-115 <VAN>
A; Residues: BMBL:229450; NID:9443862; PIDN:CAA82588.1; PID:9443863
A; Experimental source: genotype 1, N5
A; Experimental source: genotype 1, N5
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <NAT>
                                                                                                                                                                                                                                                                            R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                          polyprotein - hepatitis C virus (genotype 1, N5) (fragment)
            PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Search completed: August 7, 2003, 11:21:49
Job time : 10.5455 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                   C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                         A; Reference number: S41341
                                                                                                                                                            N;Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                        C; Accession: S41347
                                                                                                                                                                                                                                                                                                                                                                                  A; Accession: S41347
                                                                                      RESULT
            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    pecies: hepatitis C virus
A;Variety: genotype 1, N3
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41345
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
A; Residues: 1-115 <VAN>
A; Cross-references: EMBL: Z29453; NID: 9443868; PIDN: CAA82591.1; PID: 9443869
A; Experimental source: genotype 1, N8
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A,Accession: S41345
A,Molecule type: genomic RNA
A,Residues: 1-115 <VAN>
A;Cross-references: EMBL:229448; NID:g443858; PIDN:CAA82586.1; PID:g443859
                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (genotype 2, N3) (fragment)
N:Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ome polyprotein - hepatitis C virus (genotype 1, N3) (fragment)
                                                                                                                                                                                Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 138; DB 2; Length 115;
Pred. No. 2.2e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 115
                                                                                                                                                                                                                              1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Experimental source: genotype 2, N3
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Experimental source: genotype 1, N3 C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F; 1-115/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 138; DB 2;
Pred. No. 2.2e-12;
2; Mismatches 1
                                                                                                                                                                             Score 138; DB 2;
Pred. No. 2.2e-12;
2; Mismatches 1
                                                                                                                                                                                                                                                                            1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                        1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92.0%;
89.3%;
                                                                                                                                                                           Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A Accession: 841354
A; Molecule type: genomic RNA
A; Residues: 1-115 CVAN>
A; Cross-references: EMBL:229457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 92.0
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 92.0
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ontains: core protein
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

PKPQRKTERNTNRRPQDVRFSGGGQIVG 28 

ö

Gaps

ö

Length 115; Indels

92.0%; Score 138; DB 2; 89.3%; Pred. No. 2.2e-12; tive 2; Mismatches 1;

```
GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 11:05:41; Search time 4.90909 Seconds

(without alignments)

268.226 Million cell updates/sec

Title: US-09-491-146A-31
Perfect score: 150
Sequence: 1 PKPORKTERNTNRRPQDVRFSGGGIVG 28

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5
```

al number of hits satisfying chosen parameters: 127863

127863 seqs, 47026705 residues

Searched:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

		Description	59 hepatitis c			hepatitis			52 h genome po	58 h genome po	h genome	ב	ď	59 h genome po	c	38 h serine/th	0	76 herpes simp	58 oncorhynchu	73 xenopus lae	51 nicotiana t		49 escherichia					-	07 vibrio para		45 homo sapien		15 aspergillus		25 hacmonhilts
÷		Desc	P27959	001404	001403	P27960	P27961	P26663	P266	P279	P26660	P26661	P29846	000269	P26664	Q14738	028653	P28276	P16058	P21573	080361	P45978	P25549	Q97a95	88n66O	Q9ez11	Q8nws4	P17889	Q87s07	P05059	P10645	P40036	013415	09r244	200710
SOMMANDS		ជ	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVTW	POLG_HCVJT	POLG_HCV1	2A5D_HUMAN	2A5D_RABIT	IE63_HSV2H	ESR1_ONCMY	YB1_XENLA	RK4_TOBAC	SCD6_YEAST	ASLA_ECOLI	TRI_THEVO	DAPB_STAAM	DAPB_STAAU	DAPB_STAAW	IF2_BACSU	GUAA_VIBPA	CMGA_BOVIN	CMGA_HUMAN	GIP2_YEAST	AREA_ASPOR	TRP2_MOUSE	
		60	-		Н	-	٦	-			٦	-1	~		-	-	٦	-1	-	<del>, -</del>	-	-	-	٦		-	-	-	٦	٦	П	-	-	-	•
		Match Length	513	520	520	737	737	3010	3010	3011	3033	3033	3010	3010	3011	602	586	512	622	303	282	349	551	1030	240	240	240	716	517	449	457	548	998	1172	5
æ	Ouery	Match	92.0	92.0	92.0	92.0	92.0	92.0	92.0	92.0	95.0	92.0	87.3	86.7	86.7	35.7	35.0	34.7	33.3	33.0	32.0	32.0	31.3	31.3	30.7	30.7	307	30.7	30.3	30.0	30.0	30.0	30.0	30.0	6
		Score	138	138	138	138	138	138	138	138	138	138	131	130	130	53.5	52.5	52	D	49.5	48	48	47	47	46	46	46	4	45.5	45	45	45	45	45	7 7 7
	Result	Š.	-	7	e	4	ς.	· •	7	ω	σ,	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	32	66

ASMU COLL6 COCL1 AALT1 GOAIN UDAN ICI NROME RROME RROME RROM3	1 GUAA_PASMU GUAA_ECOL6 1 GUAA_ECOL1 1 GUAA_SCATT 1 GUAA_SALTI 1 LCT2_BOVIN 1 TCPL_HUMAN 1 RS2_SPICT 2 SR5_DROME 1 WNS2_EHDY2 1 WAS2_EHDY2 1 NCAP_CYMAS	523 1 GUAA_PASMU 525 1 GUAA_ECOL6 525 1 GUAA_ECOL1 525 1 GUAA_SALTI 151 1 LCT2_BOVIN 215 1 TCPL_HUANN 359 1 RSZ_SPICT 375 1 SR55_DROME 376 1 VRSZ_EHDYZ 440 1 VD49_YEAST 454 1 NCAP_CYM3	ненененене	5523 5525 5525 5526 5526 5526 5526 5526	Q9cnx8 pasteurella	P04079 escherichia	Q8z4q3 salmonella	062644 bos taurus	Q8tdr4 homo sapien	P19679 spiroplasma	P26686 drosophila	P27280 epizootic h	Q02354 saccharomyc	P18447 murine coro	P03416 murine coro
	1 GUAA_1 1 GUAA_1 1 GUAA_1 1 GUAA_1 1 TCT2_1 1 TCT2_1 1 TCT2_1 1 SR52_1 1 YNS2_1 1 YOA49_1	ненененене	5523 5525 5525 5526 5526 5526 5526 5526	299.7 299.7 299.7 299.3 299.3 299.3 299.3 299.3 299.3 299.3 299.3 299.3 299.3 299.3 299.3 299.3 299.3 299.3 299.3 299.3 299.3											

## ALIGNMENTS

POLG_HOVA2  DOLG_HOVA2  DOLG_HOVA3  DOLG_H	HCVJ2 POLG_HCVJ2 STANDARD; PRT; 513 AA. P27959; 01-AUG-1992 (Rel. 23, Created) 16-OCT-2001 (Rel. 24), Last sequence update) 16-OCT-2001 (Rel. 40, Last annotation update) Genome polyprotein [Contains: Capsid protein C (Core protein) (P22); Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GF68) (GP70) (NS1)] (Fragment). Hepatitis C virus (isolate HC-J2) (HCV). Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepatilin;	SEQUENCE FROM N.A.  SEQUENCE FROM N.A.  SEQUENCE FROM N.A.  SEQUENCE STORTS: Pubmed-1314459;  Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,  Tanaka T., Fukuda S., Tsuda F., Mishiro S.;  Tanaka T., Fukuda F., Mishiro S.;  Full-length sequence of a hepatitis C virus genome having poor  homology to reported isolates: comparative study of four distinct  genotypes.;  Yurology 188:331-341(1992).  -I- FUNCTION: THE SMALL PROTEINS NSZA, NSZB, NS4A AND NS4B ARE  HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.  NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.  NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.  NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL STORTENS:  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF	ppsid.  1.  Coat protein; Envelope protein;  REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPETIONE. CAPSID PROTEIN C OPTENTIAL). MATRIX PROTEIN (POTENTIAL). MAJOR ENVELOPE PROTEIN E (POTENTIAL).
	STANDARD; 1. 23, Crea 1. 23, Last 1. 40, Last 1. 40, Last 1. 40, Last 1. 51, Crotal 1. 51, Crotal 1. 51, Crotal 2. 51, Crotal 3. 51, Crotal 3. 51, Crotal 5. 51, Crotal 5. 51, Crotal 5. 51, Crotal 6. 51, Crotal 6. 51, Crotal 7.	A. 2: PubMed-1 11 K., Okad 12 B., Tsud 13 B., Tsud 13 C., Tsud 13 C., Tsud 13 C., Tsud 13 C., Color 14 C., Color 15 C., Color 16 C., Color 17 C., Color 18 C., Co	AA00968.1; 2522; HCV_C 2521; HCV_C 2531; HCV_N HCV_Capsid; HCV_Capsid; HCV_Capsid; HCV_Capsid; HCV_Capsid; HCV_Capsid; HCV_NS1; 1. 1 HCV_NS1;
	10012 (8) 31-1992 (Re- 11992 (Re-	VCE FROM N  T., Fuku  T.,	D10074; B Pro: IPR000 Pro: IP
	מר פין		•

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-9101900; Pubmed-1383400;

ADE K., Inchauspe G., Fujisawa K.;

ADE K., Inchauspe G., Fujisawa K.;

AGENOMIC characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";

J. Gen. Virol. 73:2725-279(1992).

-I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIDPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGNS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MNNA.
                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                       01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last sequence update)
Genome polyprotein (Contains: Capsid protein C (Core protein (P22)
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GR68) (GP70) (NS1)] (Fragment).
Hepatitis C virus (isolate HCV-476) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
   NONSTRUCTURAL PROTEIN NS1 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               protein.
REMOVED FROM CAPSID PROTEIN C BY THE
CELLULAR AMINOPEPTIDASE.
                                                                                    (POTENTIAL)
                                                                                                                (POTENTIAL)
                                           (POTENTIAL)
                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                ;
0
                                                                                                                                                                                                                  DB 1; Length 513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                       943F31E3514CDEF3 CRC64;
                           N-LINKED (GLCNAC.
                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                Score 138; DB 1
Pred. No. 8e-13;
                                                                                                                                                                                                                                                                                                                                                                            520 AA
                                                                                                                                                                                                                                             2; Mismatches
                                                                                                                                                                                                                                                                         PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002522; HCV_capsid.
InterPro; IPR002512; HCV_core.
InterPro; IPR002513; HCV_env.
InterPro; IPR002531; HCV_env.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
                                                                                                                                                                                     55704 MW;
                                                                                                                                                                                                                                                                                                                                                                                                      01-JUL-1993 (Rel. 26, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Transmembrane; Nonstructural
                                                                                                                                                                                                                  92.08;
                                                                                                                                                                                                                              ilarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D10688; BAA01530.1; -
                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
 5513
369
2233
2233
334
250
4423
4430
                                                                                                                                                       448
513
513 AA;
                                                                                                                                                                                                                              Local Similarity
es 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF01560;
                                                                                                                                                                                                                                                                                                                                                                         POLG_HCVH4
Q01404;
                                        CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
               TRANSMEM
                                                                                                                                                                     NON_TER
SEQUENCE
                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INIT_MET
                                                                                                                                                                                                                                                                                                                                                            POLG_HCVH4
                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                              RESULT 2
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                      ö
CAPSID PROTEIN C (POTENTIAL).
MATIX PROTEIN (POTENTIAL).
MAJOR ENVELODE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
BY SIMILARITY.
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-93019030; PubMed-1383400;
Abe K., Inchauspe G., Fujisawa K.;
Genomic characterization and mutation rate of hepatitis C virus
'Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
I. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus (isolate HCV-KF) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                  Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_env.
InterPro; IPR005531; HCV_env.
InterPro; IPR005531; HCV_ensi.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01559; HCV_env; 1.
Probom; PD18606; HCV_NS1; 1.
Probom; PD18606; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                                                                                                      1; Indels
                                                                                                                                                                                                                                                 AA135246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                   Score 138; DB 1;
Pred. No. 8.1e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  520 AA
                                                                                                                                                                                                                                                                                                                      2; Mismatches
                                                                                                                                                                                                                                                                                                                                                       28
                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; JQ1925; JQ1925.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                 56499 MW;
                                                                                                                                                                                                                                                                                  92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D10687; BAA01529.1; -.
                                                                                                                                                                                                                                                                                                Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=31644;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVHK
Q01403;
                                                      CHAIN
TRANSMEM
CARBOHYD
                                                                                                                                                                                                                              NON_TER
SEQUENCE
                                                                                                         CARBOHYD
                                                                                                                        CARBOHYD
                                                                                                                                            CARBOHYD
                                                                                                                                                           CARBOHYD
                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVHK
   à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                   ö
                           CAPSID PROTEIN (POTEWIAL).
MATRIX PROTEIN (POTEWIAL).
MAJOR ENVELORE PROTEIN E (POTEWIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTEWIAL).
N-LINKED (GLCNAC. .) (POTEWIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              protein); Matrix
E; Nonstructural
                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VITOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDOPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION. NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSIGNS OF TWO PROTEIN MAY DIXCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN A ND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MANA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus (isolate HC-J5) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-92230232; PubMed-1314459; Okamamoto K., Lizuka H., Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes.";
BY THE
                                                                                                                                                                                                                                                                                                                                                                 ö
PROTEIN C
                                                                                                                                                                                                                                                                                                                             Length 520;
                                                                                                                                                                                                                                                                                                           Score 138; DB 1; Length 52. Pred. No. 8.le-13;
                                                                                                                                                                                                                                                                                               MW; 1D2BD0A6FF27349B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein (Envelope protein M); Major envelope protein proteins NSI and NS2) (Fragment)
                  AMINOPEPTIDASE
FROM CAPSID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         737 AA.
                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                      CELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Interpro; IPR002522; HCV_capsid.
Interpro; IPR002521; HCV_core.
Interpro; IPR002519; HCV_core.
Interpro; IPR002513; HCV_sol.
Pfam; PF01543; HCV_capsid: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D10075; BAA00969.1; -.
                                                                                                                                                                                                                                                                                                                               92.08;
                                                                                                                                                                                                                                                                                                                                                 89.38;
                                                                                                                                                                                                                                                                                               56476
                                                                                                                                                                                                                                                                                                                                                 est Local Similarity 89.3
atches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                           520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLG_HCVJ5
ID POLG_HCVJ5
                                                                                                                                                                                                                                                          CARBOHYD
NON_TER
SEQUENCE
                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                       CARBOHYD
                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                               Query Match
INIT_MET
                                                                                                   TRANSMEM
                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                Matches
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                             CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/REZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/REZ (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-AUG-1992 (Rel. 23, Last Sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        proteins NS1 and NS2] (Fragment).

Hepailtis C virus (isolate HC-J7) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -1. FUNCTION: THE SMALL PROPEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION. NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION. 1. SUBJUNIT: THE VIRION OF THIS VIRUS IS A NICLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-92230232; PubMed-1314459; Okamento K., Lizuka H., Okameto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda F., Mishiro S.; "F., Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                    REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX PROTEIN C AND MRNA.
                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                            Length 737;
                                                            protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB.1; Length /3
Pred, No. 1.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                  3AF699D82AD501B1 CRC64;
                                                                                                  CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                         N-LINKED (GLCNAC.
                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                      (GECNAC
                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       737 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PKPORKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                             N-LINKED
N-LINKED
                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                  N-LINKED
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat pr
Transmembrane; Nonstructural protein INII_MET
                                                                                                                                                                                                                                                                                                                                                                                                  MW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last seq
                                                                                                                                                                                                                                                                                                                                                                                                                        92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Virology 188:331-341(1992)
                                                                                                                                                                                                                                                                                                                                                                                                  81207
                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity 89.3
25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLG_HCVJ7
P27961;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genotypes.
                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --
                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local S
                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                             TRANSMEM
                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                    CHAIN
                                                                                                              CHAIN
                                                                                                                           CHAIN
                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                    datches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15-5292 (Rel. 23, Created)
01-AGG-1992 (Rel. 23, Last sequence update)
15-52P-2003 (Rel. 24, Last annoctation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (MS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/halicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NSA (P4); Nonstructural protein NSAB (P27); Nonstructural protein NSAB (P56); Nonstructural protein NSAB (P56); Nonstructural protein NSAB (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (Isolate BK) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                          CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 138; DB 1; Length 737; Pred. No. 1.2e-12;
                                                                                                                                                                                                                                                                                                                                       Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                         CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC,
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                      Nonstructural protein.
                                                                                                                                                             InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002531; HCV_core.
InterPro; IPR002531; HCV_cenv.
InterPro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01540; HCV_core; 1.
Pfam; PF01550; HCV_env; 1.
Probom; PD186062; HCV_env; 1.
Probom; PD186062; HCV_env; 1.
Transmembrane; Nonstructural protei
                                                                                                                                                  EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 89.3
les 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                        INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVBK

ID POLG_HR

POLG_HR

DT 01-AUG

DT 11-AUG

DT 15-SEP

DE GENVELO

DE (FC 3.

DE (FC 3.

DE NSSB (

DE NSSB (

OS HEPATI

OS HEPATI

OX VITUSE

OX HEPATI
                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Watches
                     g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; "Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                               X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDILINE-9701508B; pubmed-8651916,
MICKPETSHAM J.A., HOSTOMSKY Z., Habuka N.,
MOOMAW E.W., Adachi T., Hostomska Z.; Hostomsky Z., Habuka N.,
The crystal structure of hepatitis C virus NS3 proteinase reveals
trypsin-like fold and a structural zinc binding site.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (RNA)(N).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                          MEDLINE-96335224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by CAMP-dependent protein kinase.";
Eur. J. Blochem. 237:611-618(1996).
MEDLINE-91140698; PubMed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                     from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M58335; AAA72945.1; -.
                                                                                                                                                          SEQUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14-JAN-98.
08-APR-98.
15-NOV-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10-APR-02.
26-JUN-00.
20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-NOV-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    interPro; IPR001410;
interPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-MAR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .nterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10UV;
80HM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1A10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1GX6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1CSJ;
1GX5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LNS3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PDB;
```

us-09-491-146a-31.rsp

```
7. 7.200.64;
7. 01-AUG-1992 (Rel. 23, Created)
7. 01-AUG-1992 (Rel. 23, Last sequence update)
7. 10-AUG-1992 (Rel. 23, Last sequence update)
7. 10-AUG-1992 (Rel. 41, Last annotation update)
7. 10-AUG-1992 (Rel. 41, Last annotation update)
8. 10-AUG-1992 (Rel. 41, Last annotation E2
8. 10-AUG-1992 (Rel. 41, Last annotation Rel. 42, Last annotation 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                {RNA}{N}.
SUBJUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE=91088550; Pubmed=2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
Japanese patients with non-A, non-B hepatitis.",
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 1; Length 3010;
Pred. No. 5.5e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                       327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPORKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Match 92.0%;
Local Similarity 89.3%;
hes 25; Conservative ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                         1688
                                                                                                                                                                                                                                                                                                                                                                                                                                              Ä
                                                                                                                                                                                                                                                                                                                                                                                                                1680
                         11120
11122
11139
11149
11162
11165
                                                                                                                                                                                                                                                                     1172
1175
1187
1189
1198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVJA
P26662;
                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
STRAND
STRAND
STRAND
STRAND
TURN
                                                                                                                      STRAND
STRAND
HELIX
TURN
TURN
STRAND
TURN
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | DECT | POLG | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                              TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 7
  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELLUIAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -LINKED (-LINKED (-LI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LINKED
                                        InterPro; IPR002868; HCV_NS5a.
InterPro; IPR002166; HCV_NS5a.
InterPro; IPR007095; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_PSvir.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01542; HCV_Core; 1.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01504; HCV_NS5a; 1.
Pfam; PF016098; Viral_RRP; 1.
Probom; PD186662; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    645
2041
2041
2240
2240
2788
2788
1047
1063
1074
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1081
1085
1087
1092
1094
1097
1103
                      InterPro; IPR001490;
InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            645
2041
2077
2240
2529
2788
1031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1093
1095
1101
1104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3D-structure, INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HELIX
STRAND
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TURN
```

ö

Gaps

œ

Fri Aug

```
of unwinding."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          [RNA] (N).
                                                                                                                                                                                                                                                                                                                                                  POLG_HCVH
P27958;
     CARBOHYD
CARBOHYD
                                                   CARBOHYD
CARBOHYD
CARBOHYD
                                     CARBOHYD
                                                                                                  CARBOHYD
                                                                                                                   CARBOHYD
                                                                                                                                  SEQUENCE
                                                                                                                                                                   Query Match
                                                                                                                                                                                    Best Loc
Matches
                                                                                                                                                                                                                                                                                                                RESULT 8
POLG_HCVH
                                                                                                                                                                                                                                                               g
     FT FT FT S
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the SWISS institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR ANINOPERTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
PROTEIN C AND MRNA. SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001650; Helicase_C.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001650; RNA_pol_DS_PS.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR001543; HCV_core; 1.
Ffam; PF01540; HCV_core; 1.
Ffam; PF01550; HCV_NS1; 1.
Ffam; PF01560; HCV_NS2; 1.
Ffam; PF01001; HCV_NS3; 1.
Ffam; PF01001; HCV_NS4a; 1.
Ffam; PF01506; HCV_NS5a; 1.
Ffam; PF01806; HCV_NS5a; 1.
Ffam; PF01806; HCV_NS5a; 1.
                                                                                                                                                                                                                                                                            HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                              InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR0004109; HCV_NS3.
InterPro; IPR001499; HCV_NS4a.
InterPro; IPR001499; HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002166; HCV_RdRP
                                                                                                                                                                             EMBL; D90208; BAA14233.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002519;
InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                            PIR; A39253; GNWVCJ.
HSSP; P26663; LJXP.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                            InterPro; IPR002522;
                                                                                                                                                                                                                                                               InterPro; IPR001410;
                                                                                                                                                                                                                                                                                               InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1863
2014
2014
347
11108
11107
12165
1316
209
209
4417
4417
4417
532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
THE PROPERTY OF THE PROPERTY AND DESCRIPTION OF THE PROPERTY A
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protean P7; Nonstructural protein NS2 (P21)
(EC 3.4.29-.); Protease/hellcase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-9915431; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonuclectide: the crystal structure provides insights into the mode
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-1187654;
Yao N., Hesson T., Cable M., Hoop Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Blol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ပ
                                     (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                           (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
 (POTENTIAL)
                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                           Length 3010;
                                                                                                                                                                                                                               Indels
                                                                                                                                                       CRC64;
                                                                                                                                                   MW; AA993794F46DB185
                                                     (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                         Score 138; DB 1;
Pred. No. 5.5e-12;
                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                       3011 AA
                                                                                                                                                                                                                               2; Mismatches
                                                                                                                                                                                                                                                                   1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                       92.08;
                                                                                                                                                   327017
                                                                                                                                                                                                             89.38;
                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Structure 6:89-100(1998)
                                     623
645
2041
2077
2240
2788
                                                                                                                                                       Ä
                                                                                                                                                                                                         Local Similarity
les 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus.
NCBI_TaxID=11108;
                                   623
645
2041
2077
2240
2788
3010
```

us-09-491-146a-31.rsp

```
(POTENTIAL)
    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 772CBB29CCD94753 CRC64;
     (GLCNAC
     N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ¥,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 327142
                                                                                                                                                                                  AA;
                                                                                                                                                                                               CARBOHYD
TUURN
STRAND
STRAND
STRAND
STRAND
CTURN
STRAND
CTURN
STRAND
CTURN
STRAND
CTURN
    CARBOHYD
CARBOHYD
CARBOHYD
     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2 THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA. PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ВУ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5B.
POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMIL DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE.
CERSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
                                                                                                                                                                                                                                                                                              TREANSFACT T04155; -
InterPro; IPR001410; DEAD.
InterPro; IPR001511; HCV_capsid.
InterPro; IPR002511; HCV_capsid.
InterPro; IPR002511; HCV_capsid.
InterPro; IPR002511; HCV_capsid.
InterPro; IPR002519; HCV_mS1.
InterPro; IPR001409; HCV_NS2.
InterPro; IPR001409; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR00166; HCV_NS4.
InterPro; IPR001650; Hellcase_C.
InterPro; IPR001650; Hellcase_C.
InterPro; IPR001650; HCV_NS4.
InterPro; IPR001650; HCV_NS4.
InterPro; IPR001650; HCV_NS3.
InterPro; IPR001650; HCV_NS4.
IPfam; PF01539; HCV_copsid; I.
Pfam; PF01539; HCV_NS3; I.
Pfam; PF01506; HCV_NS3; I.
Pfam; PF01506; HCV_NS3; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS4; I.
                                                                                                                                                                                                      EMBL; M67463; AAA45534.1; -. PIR; A36814; GNWVCH. PDB: 1HEI; 25-NVY-98. PDB: 1A1V: 16-FEB-99. PDB: 1A1R: 17-JUN-98. MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  191
383
746
809
1026
1171
1197
2301
3011
1083
11083
11167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   192
384
384
747
1027
11712
11712
12973
1185
11003
11165
11230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
```

```
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                            ö
                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 11, Last annoctation update)
01-AUG-1992 (Rel. 11, Last annoctation update)
03-FREL 11, Last annoctation update)
04-Ond (Contains: Capala)
05-Ond (Capala)
05-Ond
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEDUENCE FROM N.A.

MEDLINE-9204440; PubMed-1658196;

MEDLINE-92044440; PubMed-1658196;

Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,

Machida A., Miyakawa Y., Mayuul M.;

Mucleotida A., Miyakawa Y., Mayuul M.;

"Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";

J. Gen. Virol. 72:2697-2704(1991).

-I. FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLOTION.

-I. CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.

CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate
                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
Score 138; DB 1; Length 3011;
Pred. No. 5.5e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                             3033 AA
                                                                                                           1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                           HCV_caps1d.
HCV_core.
HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS3.
92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D00944; BAA00792.1; -.. PIR; JQ1303; JQ1303.
HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410; DEAD.
                          Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR000745;
IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                    POLG_HCVJ6
P26660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                       POLG_HCVJ6
ID POLG_HC
AC P16660,
DT 01-AUG-
DT 28-FEB-
DE Genome
DE Genome
DE GROWE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 <u>-</u>
                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                               g
```

```
ö
                                                                                                                                                                                                                                                                                         Pfam; PF010u1,

R Pfam; PF010u1,

R Pfam; PF00271; helicase_C; 1.

DR Pfam; PF00998; Vital_RRP; 1.

DR Probom; PD186062; HCV_NRP; 1.

DR SMART; SM0487; DEXDC; 1.

KW Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;

KW Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding;

Transmembrane; Nonstructural protein; Hydrolase; Serine protease

Transmembrane; Nonstructural protein; Hydrolase; Serine Protein; Mydrolase; Mydrolase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE,HELICARS NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BY SIMILARITY).
(BY SIMILARITY).
(BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5.6e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 138; DB 1;
Pred. No. 5.6e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 10
POLG HCVJ8
TD POLG-HCVJ8 STANDARD, PRT; 3033 AA. AC. BC561;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ...crpro; IPRO01650; Hellcase_C.
InterPro; IPRO07095; RNA_pol_DS_PS.
InterPro; IPRO07095; RNA_pol_DS_PS.
InterPro; IPRO07094; RNA_pol_PSVir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01539; HCV_core; 1.
Pfam; PF01560; HCV_cory; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        101
2038
2038
329165 MW, F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                         Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01538; HCV_NS2; 1.
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5b; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     733
1010
1619
1866
2017
3033
369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   627
649
1091
2038
2038
201
2811
2833 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRANSMEN
```

```
ARY REPARTMENT OF THE PROPERTY OF THE PROPERTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
28-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21) (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS48 (P27); Nonstructural protein NS58 (P27); Nonstructural protein NS58 (P27); Nonstructural protein NS58 (P50); Nonstructural protein Hepatitis C virus (1solate HC-38) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-92230232; PubMed-1314459; Okamoto H., Lizuka H., Tanaka T., Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; **Full-length sequence of a hepatitis C virus genome having poor penology to reported isolates: comparative study of four distinct genotypes.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VITOLOGY 188:331-341(1992).

--- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

--- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 postition, Cys or Thr in P1 and Ser or Ala in P1'.

--- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (RNA)(N).
SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROFIEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; TPR001490; HCV_NS4b.
InterPro; IPR002668; HCV_NS5a.
InterPro; IPR002668; HCV_NS5a.
InterPro; IPR007165; RNA_DOL_DS.PS.
InterPro; IPR007095; RNA_DOL_PSvir.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01560; HCV_Core; 1.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01506; HCV_NS2; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS5a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D10988; BAA01761.1; -. PIR; A40250; GNNVJ8. HSSP; P27958; JHEI. MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Interpro; IPR002519; H
Interpro; IPR002511; H
Interpro; IPR002518; H
Interpro; IPR004109; H
Interpro; IPR00145; H
Interpro; IPR00145; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11115;
                                                                                                                                                                                                                                                                                                    Hepacivirus.
```

```
ö
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
W Core protein; Coat protein; Brvelope protein; Helicase; APP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
T INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
T CHAIN 116 191 RAPRIN PROTEIN (POTENTIAL).
T CHAIN 192 383 MAPRIX PROTEIN (POTENTIAL).
T CHAIN 194 1010 MADIA PROTEIN NSI (POTENTIAL).
T CHAIN 1011 1619 PROTEASE/HELICASE NSI (POTENTIAL).
T CHAIN 1021 166 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
T CHAIN 1620 1666 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1640 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 2018 3033 RAA-DIRECTED RNA POLYMERASE (POTENTIAL).
T TRANSMEM 347 369 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-APR-1993 (Rel. 25, Last annotation update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (NSI); Protein El (GP35); Envelope glycoprotein El (GP68) (GP70) (NSI); Protein P7; Nonstructural protein NS2 (P21)
16-SEP-2003 (NSI); Protein P7; Nonstructural protein NS4 (P4); Nonstructural protein NS4 (P65); Nonstructural protein NS5 (P65); Nonstruc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INKED (GLCNAC. .) (POTENTIAL)
1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-92230206; Pubmed-1314449;
Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB 1; Length 3033;
Pred. No. 5.6e-12;
2; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (Rel. 25, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
les 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVTW
P29846;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loc
Matches
```

(BY SIMILARITY).
(BY SIMILARITY).
(BY SIMILARITY).

SYSTEM (SYSTEM (SYSTEM (

(POTENTIAL). (POTENTIAL). (POTENTIAL).

(GLCNAC.

(POTENTIAL)

(GLCNAC

```
ATP (POTENTIAL)
DECH BOX.
                            RELAY
       CHARGE RELAY
                 CHARGE RELAY
CHARGE RELAY
                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                           2 KPORKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                                                                                                                                                                                                                                                         87.3%;
88.9%;
                                                                                                                                                                                                                                                                                                                                  327047
                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 88.9
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=31642;
                                                                                                                                                                                                                                           623
645
2041
2047
2240
2529
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (RNA) (N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCVJT
                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                   CARBOHYD
                                                                                                       CARBOHYD
                                                                                                                     CARBOHYD
                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 12
POLG_HCVJT
     à
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropaen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions along as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR ANINOPERIDAGE.
CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
POTENTIAL).
POTENTIAL).
                                FUNCTION: THE SMALL PROFEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RRA REPLICATION.

CATALYTIC ACTIVITY: HYDROLYSIS OF four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.

CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).

SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROFEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN CAND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
"The Talwanese hepatitis C virus genome: sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA."; Virology 188:102-113(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Envelope protein; Helicase; ATP-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REMOVED FROM CAPSID PROTEIN C BY
                                                                                                                                                                                         SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                         or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001490; HCV_NS4b.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR001066; HCV_RGAP.
InterPro; IPR007095; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_PSvir.
Pfam; PF01543; HCV_Core; 1.
Pfam; PF01543; HCV_Core; 1.
                                                                                                                                                                                                                                                                                                                  EMBL; M84754; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                          HCV_capsid
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; hellcase_C; 1.
Pfam; PF00909; VIral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR000745; HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002531; HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Core protein; Coat protein;
                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410;
InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                          1N64; 25-FEB-03
1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                               A40244; GNWVTW.
                                                                                                                                                                                                                                                                                                                                                                   MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01560;
PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF01006;
Pfam; PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
```

```
ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   000269;
01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
16-SEP-2003 (Rel. 25, Last sequence update)
16-SEP-2003 (Rel. 42, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein E2 (GP68) (GP70) (NS1); Protein E1 (GP32) (GP55); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein Rel (GP32) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5B (PC2); Nonstructural protein NS5A (PS6); Nonstructural protein NS5B (PC5) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-J1) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VILUE RES. 23:99-53(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 postition, Cys or Thr in P1 and Ser or Ala in P1'.

-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE-9295714; Pubmed=1318627;
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Nakazawa T., Hijikata M., Ishimura Y., Shimotohno K.;
"Modecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among
                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
(POTENTIAL)
                                                                                                                     (POTENTIAL)
                                                                         (POTENTIAL)
                                                                                                (POTENTIAL)
                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                   (POTENTIAL
                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                  DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                             1; Indels
                                                                                                                                                                                                                                                                                                                  AAD267D55CDFE215
                                             (GLCNAC.
                                                                                                                                                                   (GLCNAC.
                                                                                                                  (GLCNAC.
                         GLCNAC
                                                                                                (GLCNAC.
GLCNAC
                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                      6e-11;
                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
                                                                                                                                                                                                                                                                                                                                                               Score 131;
Pred. No. 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6 KPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                28
                                                                                         N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                        N-LINKED
N-LINKED
N-LINKED
```

(POTENTIAL)

POTENTIAL POTENTIAL

ö

Gaps

ö

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@1sb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-505 (Rel. 23, Created)
01-505 (Rel. 23, Last sequence update)
01-507 (Rel. 42, Last annotation update)
62003 (Rel. 42, Last annotation update)
6300 (GP68) (Rel. 42, Last annotation update)
6300 (GP68) (GN1); Protein F1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein F2 Nonstructural protein NS2 (P21)
630 (GP70) (NS1); Protein F2 NS3 (P70) (Repacivirin)
630 (GP70); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
64 (P40) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
65 (P60) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (RNA)(N).

SUBURIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                   (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                  (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 86.7%; Score 130; DB 1; Length 3010;
85.7%; Pred. No. 8.5e-11;
1ve 2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                               CRC64;
                                                                                                                                                                                                                                                                                                                                                                                               MW; 94A1C77435D642BB
                                                                                                                                                                   (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                  (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                             N-LINKED
N-L
                                              448
532
532
532
540
56
56
56
576
576
523
5207
2207
2240
2529
277
2240
2529
278
3 326573 M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity 85.7
nes 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11104;
                                                                                                               540
556
576
623
6623
2041
2077
2077
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCV1
                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                       CARBOHYD
                                                                                                                  CARBOHYD
                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P26664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 13
POLG_HCV1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DATE TO THE TOTAL 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ç
                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                            SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELLULAR AMINOPERTIDASE.
CASID PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOR PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROFEASE/HELICASE NS3 (POTENTIAL).
PROMSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LUCERPIO, IPRODI490; HCV_NS4a.
LUCERPIO; IPROD2868; HCV_NS4b.
INTERPIO; IPROD2166; HCV_NS5a.
INTERPIO; IPROD7095; RNA_POI_DS_PS.
INTERPIO; IPROD7095; RNA_POI_DS_PS.
INTERPIO; IPROD7095; RNA_POI_PSVIT.
Pfam; PFOI543; HCV_core; I.
Pfam; PFOI542; HCV_core; I.
Pfam; PFOI559; HCV_core; I.
Fam; PFOI560; HCV_NS1'
Fam; PFOI560; HCV_NS1'
Fam; PFOI560; HCV_NS1'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHARGE
ATP (PO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF01506; HCV_NSAs; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D11168; BAA01943.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF01539; HCV_env; 1.
PF01560; HCV_NS1; 1.
PF01538; HCV_NS2; 1.
PF02907; HCV_NS3; 1.
PF01006; HCV_NS4a; 1.
PF01001; HCV_NS4b; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1006
11615
12862
12003
13010
11003
11003
11003
1234
1319
1319
1319
1309
1309
1309
1309
                                                                                                                                                                                                                                                                                                                                                                                                                PIR; A45573; A45573.
PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; S29.001; -. MEROPS; U39.001; -. InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002531;
InterPro; IPR002518;
InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1616
1863
2014
347
1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
```

```
TISSUENCE FROM N.A. (ISOFORMS DELTA-1 AND DELTA-2).

TISSUE-Colon, Eye, and Kidney;

MEDLINE-218825; PubMed-1247792;

Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,

Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Hopkins R.F., Jordan H., Moore T., Mang J., Hsich F.,

A platchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

A platchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

A platchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

A platchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

Raba S.S., McEwan P.J., McKernan K.J., Malek J.A., Guneratene P.H.,

A Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

Rabhy J., Helton B., Ketteman M., Madan A., Kodrigues S., Sanchez A.,

Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

Rockiguez A.C., Grimwood J., Schmutz J., Myers R.M.,
                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TISSUE-Fetal brain; (TOTO) DESTRUCTORY STRUCTORY STRUCTO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. (ISOFORM DELTA-2), AND SEQUENCE OF 501-508; 550-559; 573-580 AND 584-601 (DELTA-1).

TISSUE-Bone marrow, and Brain cortex; MEDLINE-96159032; PubMed-64566219;

Tanabe O., Nagase T., Murakami T., Nozaki H., Usui H., Nishito Y., Hayashi H., Kagamiyama H., Takeda M.; Molecular cloning of a 74-Kba regulatory subunit (B'' or delta) of human protein phosphatase 2A."; FEBS Lett. 379:107-111(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       014738; 000494; 000696; 015171; 004.

30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
Serine/threonine protein phosphatase 2A, 56 kDa regulatory subunit, delta isoform (PPAA, B subunit, B' delta isoform) (PPAA, B subunit, PSG delta isoform) (PPAA, B subunit, PRG delta isoform) (PPAA, B subunit, RS delta isoform).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens (Human).
Pustaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE-Brain cortex;
MEDLINE-97324098; PubMed-9180267;
MEDLINE-97324098; PubMed-9180267;
Tanabe O., Gomez G.A., Nishito Y., Usul H., Takeda M.;
"Molecular heterogeneity of the cDNA encoding a 74-kDa regulatory subunit (B', or delta) of human protein phosphatase 2A.";
FEBS Lett. 408:52-56(1997).
                                                                                                                                                                      ö
                                                                                             Length 3011;
                3011 AA; 327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                                                     Score 130; DB L; L
Pred. No. 8.5e-11;
3; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A. (ISOFORMS DELTA-1 AND DELTA-3)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       602 AA
                                                                                                                                                                                                                                           1 PKPORKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                    PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. (ISOFORM DELTA-1).
                                                                                     Ouery Match 86.7%;
Best Local Similarity 82.1%;
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2A5D_HUMAN
                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PPP2R5D
                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
2A5D_HUMAN
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CELLULAR AMINOPERTIDASE.
CARSID PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Hellcase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (BY SIMILARITY).
(BY SIMILARITY).
(BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
CHARGE RELAY SYSTEM ()
CHARGE RELAY SYSTEM ()
CHARGE RELAY SYSTEM ()
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                  INTERPROPTION DEAD.
INTERP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF01543; HCV_capsid; 1. Pfam; PF01542; HCV_capsid; 1. Pfam; PF01542; HCV_capsid; 1. Pfam; PF01542; HCV_capsid; 1. Pfam; PF01540; HCV_NS1; 1. Pfam; PF01506; HCV_NS2; 1. Pfam; PF01006; HCV_NS4; 1. Pfam; PF01001; HCV_NS4; 1. Pfam; PF00091; HCV_NS5a; 1. Pfam; PF00998; Viral_RdRP; 1. Pr000m; PD186062; HCV_NS1; 1. Pr0000m; PD186062; PD186062; 1. Pr0000m; PD186062; PD186062; 1. Pr0000m; PD186062; PD18606
PIR; A39166; GNWVC3.
PDB; 1A1V; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
MEROPS; S29. 001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
```

)

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation for European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TAS.
                                                                                                                                                                                                                                                COMPARTMENT:
COMPARTMENT:
COMPARTMENT:
SUBURIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
COMPOSED OF A 36 KDB CATALYTIC SUBURIT C) AND A 65 KDB
CONSTANT REGULATORY SUBURIT (PR65 OR SUBURIT A), THAT ASSOCIATE
WITH A VARIETY OF REGULATORY SUBURITS. PROTEINS THAT ASSOCIATE
WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBURITS
B (THE R2/APASS) R3/B<sup>3</sup>, 'PR72/PR130/PR59 AND R5/B'/B56
FAMILIES), THE 48 KDB VARIABLE REGULATORY SUBURIT, VIRAL PROTEINS,
SUBCELLUAR LOCATION: CYTOPLASMIC AND NUCLEAR IN INTERPHASE,
NUCLEAR DURING MITOSIS.
ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIGNAL (POTENTIAL).
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
-!- FUNCTION: THE B REGULATOR YEBURT MODULATE SUBSTRATE SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE SPECIFICITY: ISOFORM DELTA-2 IS WIDELY EXPRESSED. ISOFORM DELTA-1 IS HIGHLY EXPRESSED IN DELTA-1 IS HIGHLY EXPRESSED IN BRAIN.
INDUCTION: BY RETINICE ACID; IN WEDROBLASTOMA CELL LINES.
PTM: AT LEAST ISOFORM DELTA-1 IS PHOSPHORYLATED ON SERINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY: Belongs to the phosphatase 2A regulatory subunit B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8 X 2 AA APPROXIMATE TANDEM REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOLG407. -. GO:0005634; C:nucleus; TAS. GO:0005634; C:nucleus; TAS. GO:0008601; F:protein phosphatase type 2A, regulator acti. GO:0007399; P:neurogenesis; TAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nuclear protein; Phosphorylation; Alternative splicing; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SH3 BINDING, CLASS I (POTENTIAL)
NUCLEAR LOCALIZATION SIGNAL (POT
Missing (in isoform Delta-3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /FTId-VSP_005110.
Missing (in isoform Delta-2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FISF71AF4E565387 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Event-Alternative splicing; Named isoforms-3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IsoId-Q14738-2; Sequence-VSP_005111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IsoId-Q14738-1; Sequence-Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AB000634; BAA20381.1; -. AB000635; BAA20382.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       D78360; BAA11372.1; -. BC001095; AAH01095.1; -. BC001175; AAH01175.1; -. BC010692; AAH10692.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             69991 MM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; L76702; AAB69751.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL, BC010692, AAH10692.1
PIR, S68686, S68686.
Genew, HGNC:9312, PPP2R5D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002554; B56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             565
116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002554
Pfam; PF01603; B56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      602 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Multigene family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Name=Delta-2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Name-Delta-3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Name=Delta-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESIDUES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MIM; 601646;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3K; Q14738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SECUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
```

DB 1; Length 602;

No. 3.1;

Score 53.5; Pred. No. 3.

35.7%; 31.6%;

Best Local Similarity

Query Match

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.lab-sib.ch/announce/or send an email to license@lsb-sib.ch).
   ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Multigene family; Nuclear protein; Repeat.
DOMAIN 21 36 8 X 2 AA APPROXIMATE TANDEM REPEATS OF Q-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-96161994; PubMed-8576224;
CSortos C., Zolhierowicz S., Bako E., Durbin S.D., Depaoli-Roach A.A.;
"High complexity in the expression of the B' subunit of protein
phosphatase 2A0. Evidence for the existence of at least seven novel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Soforms.*;
J. BLOL. Chem. 271:2578-2588(1996).
J. BLOL. Chem. 271:2578-2588(1996).
-I- FUNCTION: THE B REGULATORY SUBBNIT MIGHT MODULATE SUBSTRATE SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SH3 BINDING, CLASS I (POTENTIAL).
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
E149A309CDDA7435 CRC64;
   Gaps
                                                                                                                                                                                                                                                             DASD_RABIT STANDARD; PRT; 586 AA.

028653; 028655;
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
Serine/Threorine protein phosphatase 2A, 56 KDa regulatory subunit,
Beform (PP2A, B subunit, B' delta isoform) (PP2A, B subunit,
B56 delta isoform) (PP2A, B subunit, PR61 delta isoform) (PP2A, B
subunit, R5 delta isoform) (PP2A, B subunit, B'-gamma).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Oryccolagus cuniculus (Rabbit).
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Lagomorpha, Leporidae, Oryctolagus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -!- CAUTION: NOMENCLATURE USED IN REF.1 REFERS TO PP2A B SUBUNIT B' GAMMA ISOFORM, WHICH IS CITED AS PP2A B SUBUNIT DELTA-PR61 ISOFORM IN LATER PUBLICATIONS.
   11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1; Length 586;
   4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-New Zealand; TISSUE-Brain, and Skeletal muscle;
                                                          27
                                                                                                                  81
                                                                                            |:|| ::: ::|:|| :
PQPQAQSQPPSSNKRPSNSTPPPTQLSKIKYSGGPQIV
                                                          PKPORKTE-RNTNRRPQD------VRFSGGGQIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 52.5; Di
Pred. No. 4.3;
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  35.0%; Score 31.6%; Pred.
11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          68090 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; U38193; AAC48532.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; U38195; AAC48534.1; -. InterPro; IPR002554; B56. Pfam; PF01603; B56; 1.
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               507
532
586 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=9986;
12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PPP2R5D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                        2A5D_RABIT
Matches
                                                                                                                                                                                                         RESULT 15
                                                                                                                                                                                                                                                                                                SO THE TANK AND DAY OF THE
                                                                                                            g
                                                          à
```

12; Conservative 11; Mismatches 4; Indels 11; Gaps Matches

ó

q

Search completed: August 7, 2003, 11:20:04 Job time : 5.90909 secs

œ

Fri Aug

Run on:

```
hepatitis
hepatitis
                                                                                                                                            hepatitis
hepatitis
                                                                                                                                                                              hepatitis
hepatitis
                                                                                                                                                                                                        hepatitis
hepatitis
                                                                                                                                                                                                                                            hepatitis
hepatitis
                                              hepatitis
                                                                      hepatitis
                                                                                                 hepatitis
                                                                                                                 hepatitis
                                                                                                                                                              hepatitis
                                                                                 hepatitis
                                                                                          hepatitis
                                                                                                           hepatitis
                                                                                                                          hepatitis
                                                                                                                                    hepatitis
                                                                                                                                                                       hepatitis
                                                                                                                                                                                                 hepatitis
                                                                                                                                                                                                                          hepatitis
                                                                                                                                                                                                                                    hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                K
                                                                                                                                                                                                                                                                                                                                                                                                                                               Tokita H., Kaufmann G.R., Matsubayashi M., Okuda I., Tanaka T.,
Harada H., Mukalde M., Suzuki K., Cooper D.A.;
"Heparitis C Virus Core Mutations Reduce the Sensitivity of a
"Heparitis C Virus Core Mutations Reduce the Sensitivity of a
Fluorescence Enzyme Immunoassay."
J. Clin. Microbiol. 38:3450-3452(2000).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY.
ILPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
EMBL, AB039875; BAB12423.1;
--
InterPro; IPR002222; HCV_capsid.
Pfam; PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
0668683
0668706
068707
068707
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
081750
                                                                                                                                                                                      089963
089452
099452
099454
089959
099460
                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                          103 AA
                                                                                                                                          080P64
080P55
080P60
080P57
090T58
080P63
                                                                              080P70
080P54
090T57
080P52
                                                                                                                                                                                                                                                                                                                                           Created)
Q68683
Q68706
Q688707
Q68707
Q68685
Q8JYSO
Q8JYSI
Q81264
                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-20440673; PubMed-10970401;
                                                                                                                                                                                                                                                                                                                                                                     Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                          01-MAR-2001 (TrEMBLrel. 16, 01-MAR-2001 (TrEMBLrel. 16, 01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
[1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein.
NON_TER
NON_TER 1
                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-HC-S1
                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                                                                                                                                  09E977
                                                                                                                                                                                                                                                                                                                         09E977
                                                                                                                                                                                                                                                                                                       RESULT 1
Q9E977
                                                                                                                                                                                                                                                                                                                                  hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hepatitis
hepatitis
                                                                    (without alignments)
281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                             August 7, 2003, 11:05:41; Search time 25.6364 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09e977
       5.1.6
Compugen Ltd.
                                                                                                                                                                            tal number of hits satisfying chosen parameters:
                                                                                                                1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                            830525 seqs, 258052604 residues
        GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                                           summaries
                                           protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q81481
Q8V847
Q68137
Q68311
Q68306
Q6309
Q8JYR9
Q8JYR8
Q8JYR7
Q8JYR7
                                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                      sp_unclassified:*
                                                                                                                                                                                                                                                                                               sp_human:*
sp_invertebrate:*
                                                                                                                                                                                                                                                                                                                                                                             sp_vertebrate:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ü
                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 su
                                                                                                                                                                                                                                                                                                                                                                                               sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                              sp_organelle:*
sp_phage:*
                                                                                                                                                                                              length: 0
length: 2000000000
                                                                                                                                                                                                                                                                     sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                 sp_plant:*
sp_rodent:*
sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                sp_archeap:*
                                                                                               US-09-491-146A-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length DB
                                                                                                                                                                                                                                                                                       sp_fungi:*
                                                                                                                                                                                                                                                            SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sed
                                                                                                         Perfect score:
                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score
                                            protein
                                                                                                                                                                                              2000
                                                                                                                Sequence:
                                                                                                                                                             Searched:
                                                                                                                                                                                                                                                            Database
                                                                                                                                                                                              Minimum Maximum Maximum Maximum
```

ö

Gaps

ö

Length 103; Indels

11644 MW; C2BA6559BDDDFCBC CRC64;

103

103 AA;

SEQUENCE Query Match Local

94.0%; Score 141; DB 12; 92.9%; Pred. No. 2.6e-13; 11ve 1; Mismatches 1;

1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28 

ò

hepatitis hepatitis

Q68712 Q68682 Q68708 Q68679

Result Š. 26; Conservative

Similarity

Gaps

ö

1; Mismatches

Length 191;

Fri Aug

```
191
20902 MW; 43A1B0CEFB44A0F9 CRC64;
                                                                                    94.0%; Score 141; DB 12; 92.9%; Pred. No. 5e-13;
                                                                                                                              26; Conservative
    Polyprotein.
NON_TER 191
SEQUENCE 191 AA;
                                                                                                            Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11103;
                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                    068137
                                                                                                                              Matches
                                                                                                                                                                                                                                                                              RESULT
Q68137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
068311
    FT SO
                                                                                                                                                                          à
                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                Okamoto H., Crovatto M., Pozzato G., Feray C., Brechot C., Mishiro S.; "Molecular and clinical characteristics of the hepatitis C virus genotype '2c' found in Italians in Italy and France."; Int. Hepatol. Commun. 3:161-165(1995).

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROFEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN C AND MINA (BY SIMILARITY).

FEMBL; D31986; BAA06754.1; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL, AB061932; BAB83271.1; -
InterPro; IPR00252; HCV_capsid.
InterPro; IPR004531; HCV_capsid.
Pfam; PF01343; HCV_capsid.
Pfam; PF01542; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kato N.;
"Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 128;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       128 128
128 AA; 14591 MW; 0F5F0083F8ED0B5A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         08v847;
01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                              01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-0CT-2002 (TrEMBLrel. 22, Last annotation update)
Core (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 141; DB 12;
Pred. No. 3.3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Mismatches
  PKPQRKTKRNTNRRPQDVRFPGGGGIVG 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26; Conservative.
                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                    virus
                                                                                                                                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                STRAIN-GIUS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
NON_TER
                                                                                                                                                                                                                                      Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                lesions.
                                                                                                        081481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8V847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                           RESULT 2 (081481 11D (081481 AC (081481 AC (08144 AC (08144 AC (0814 AC (08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 3
Q8V847
ID Q8V8
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STT WE BER BER SCCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-SA6;
MEDIINE-92279243; PubMed-1317578;
Bukh J., Purcell R.H., Miller R.H.;
Sequence analysis of the 5' noncoding region of hepatitis C virus.";
Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92.7%; Score 139; DB 12: Length 191; 89.3%; Pred. No. 1e-12; 1ve 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-94336721; PubMed-8058787;
Bukh J., Purcell R.H., Miller R.H.;
"Sequence analysis of the core gene of 14 hepatitis C virus
genotypes.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
NON_TER 191 191 .
SEQUENCE 191 AA; 20682 MW; 0EB8000CF1F902EB CRC64;
                                                                                                                                                            01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Proc. Natl. Acad. Sci. U.S.A. 91:8239-8243(1994).
                                                                                                                            PRT; 191 AA
                   5 PKPQRKTKRNTNRRPQDVRFPGGGQIVG 32
1 PKPQRKTERNINRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genome polyprotein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 89.3 Matches 25; Conservative
                                                                                                                            PRELIMINARY;
```

Fri Aug

```
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                     Songsivilai S., Kanistanon D., Kunkitti R.; "Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                       Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GIXCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Shustov A.V., Gavrilova I.V., Netesov S.V.;
Genetic variability of hepstitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MINA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 12; Length 60;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABDC4C21ED236AAA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                B5691CBE8F26F2F1 CRC64;
                                                                                                                                        (TrEMBLrel. 01, Last sequence update) (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB 1;
Pred. No. 3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 138; DB 13
Pred. No. 4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                60 AA.
                                                                                         Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2; Mismatches
                                                                                       46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28
PRPORKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPORKTERNTNRRPQDVRFSGGGGIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AF50616; AAM33392.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-OCT-2002 (TrEMBLrel. 22, L. 01-MAR-2003 (TrEMBLrel. 23, L. Genome polyprotein (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               60
6686 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           46 46
46 AA; 5129 MW;
                                                                                                                       01-NOV-1996 (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     U23747; AAA65054.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 89.3
Matches 25; Conservative
                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               60
AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity
                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
                                                                                                                                                                                                                                                NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                      STRAIN-HCV-BB10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-RIG306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein.
NON_TER 46
SEQUENCE 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
                                                                                                                                         01-NOV-1996
01-OCT-2002
                                                                                                                                                                                                                               Hepacivirus
'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                   068309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q8JYR9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                    RESULT 7
Q68309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OBJYR9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         virus ";
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBGNAT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GITCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN. C AND MRNA (BY SIMILARITY).

EMBL; U23744; AAA65051.1;
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-CT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. BY MIND MINN (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-HCV-BB5;
Songsivila1 S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thal isolates of hepatitis
                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                             Songsivilal S., Kanistanon D., Kunkitti R.; Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 138; DB 12; Length 45; Pred. No. 2.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 12; Length 45;
Pred. No. 2.9e-13;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CC527167096AAA81 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CC4C21ED236AAA81 CRC64;
                          01-NOV-1996 (TTEMBLrel. 01, Created)
01-NOV-1996 (TTEMBLrel. 01, Last sequence update)
01-CGT-2002 (TTEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
45 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        45 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Created)
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    45 45
45 AA; 5015 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         45 AA; 5030 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         89.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                              SEQUENCE FROM N.A. STRAIN-HCV-BB37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                       NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         45
                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
NON_TER
```

SEQUENCE

SO FT

Query Match

ö 셤 Hepacivirus

908890

NON\_TER SEQUENCE

Query Match

Matches

à

ö

Gaps

ö

K

K

```
STRAIN-RIG288;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-RIG289;
                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
                                                                                                                                                        01-0CT-2002
01-0CT-2002
                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                   Q8JYR7
                                                                                                                                              Q8JYR7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QBJYR6
                                                                                                             RESULT 11
Q8JYR7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 12
OBJYR6
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                              ò
                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                     ö
   Gaps
                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                        Shustov A.V., Gavrilova I.V., Netesov S.V.;
Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A AND GLYCORROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MANA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DBBJ databases.
'-- SUBMUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MANA (BY SIMILARITY).

EMBL, AF506620; AAM33396.1; --
InterPro; IPR002522; HCV_capsid.
Pfan, PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                     ö
   ö
                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB 12; Length 61; Pred. No. 4.1e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                    1; Indels
                                                                                                                                                                                                                                                                                                                                                              61
6897 MW; EC656DC79E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EC656DC79E8F26F2 CRC64;
                                                                                                                              01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-0CT-2002 (TrEMBLrel. 22, Created)
01-0CT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
  Τ.
                                                                                                 61 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ¥
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                    2; Mismatches
                              1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                     1 PKPQRKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                      Created)
                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                   EMBL; AF506617; AAM33393.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                    Genome polyprotein (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genome polyprotein (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                             92.0%;
89.3%;
                                                                                                                     (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
6897 MW;
25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                             61
61 AA;
                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 AA;
                                                                                                                                                                                               NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11103;
                                                                                                                                                                                                                              STRAIN-RIG286;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-RIG290;
                                                                                                                     -OCT-2002
                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
NON_TER
                                                                                                                                                               Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QBJYR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
Q8JYR5
TO ORJYR
Matches
                                                                          RESULT 9
Q8JYR8
                     ò
                                         셤
                                                                                                          OC OC DE DIA
                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A MAD GLYCOPROFIEN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MANA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Shustov A.V., Gavrilova I.V., Netesov S.V.;
Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDDJ databases
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE SWIZEDE CONSIGTS OF TWO PROTEINS:
PROTEIN M AND GLYCORPOTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AFSOGG18; ARM33394.1;
-1
InterPro: IPRO2522; HCV_capsid: 1.
                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12; Length 61;
                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EC657F706E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
              ; DB 12;
4.1e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 12;
Pred. No. 4.1e-13;
           Score 138; DB Pred. No. 4.1e-2; Mismatches
                                                                                                                                                                                                                                                               61 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
                                                                                               1 PKPORKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                         28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPORKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                                                                                                                                                                                                    (TrEMBLrel. 22, Created)
(TrEMBLrel. 22, Last seq
(TrEMBLrel. 23, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL, AF506619; AAM33395.1; ...
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genome polyprotein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                          Genome polyprotein (Fragment)
Hepatitis C virus.
         92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 61
61 AA; 6869 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-0CT-2002 (TrEMBLrel, 22, 01-0CT-2002 (TrEMBLrel, 22, 01-MAR-2003 (TrEMBLrel, 23,
Ouery Match
Best Local Similarity 89.3 Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25; Conservative
                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                 01-MAR-2003 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11103;
```

```
902890
                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 15
Q68708
                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DDT THE STAND TH
                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
STREET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A. MEDLINE-96118171; PubMed-8578855; Stuyver L., Wyseur A., van Arnhem W., Lunel F., Laurent-Puig P., Pawlotsky J.-M., Kleter B., Bassit L., Nkengasong J., van Doorn L.-J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus genotyping by means of 5'-UR/core line probe assays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Stuyver L., Wyseur A., van Arnhem W., Lunel F., Laurent-Puig P.,
Pawlotsky J.-M., Kleter B., Bassit L., Nkengasong J., van Doorn L.-J.,
                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus type lb.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 138; DB 12; Length 74; Pred. No. 5e-13;
                                                                          Length 62;
                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 74 AA; 8470 MW; 2AE2CF5ED0BD10F3 CRC64;
                                    EEEC656DC79E8F26 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-CT-2002 (TrEMBLrel. 22, Last annotation update) Core protein (Genome Polyprotein) (Fragment). Hepatitis C virus type 1b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ;
                                                                  Score 138; DB 12;
Pred. No. 4.2e-13;
2; Mismatches 1:
                                                                                                                                                                                                                                                                                       74 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         74 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                      1 PKPORKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                             (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE-96118171; Pubmed-8578855;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.0%;
89.3%;
                  62
7053 MW;
                                                                        92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 92.0
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                            Local Similarity 89.3
nes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                  62 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=31647;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=31647;
                  62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Maertens G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
Polyprotein.
                                                                                                                                                                                                                                                                                                                           01-NOV-1996
01-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                  NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Maertens
                                                                          Query Match
                                                                                                                                                                                                                                                                                   068712
068712;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       068682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 14
Q68682
                                                                                                                                                                                                                                               RESULT 13
Q68712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
FT SO
                                                                                                                                                    ò
                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                         92555
```

```
SEGUENCE FROM N.A.
MEDILINE-96118111; Pubmed-8578855;
Stuyver L., Wyseur A., van Arnhem W., Lunel F., Laurent-Puig P.,
Pawlotsky J.-M., Kleter B., Bassit L., Nkengasong J., van Doorn L.-J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Maertens G.; "Hepatitis C virus genotyping by means of 5'-UR/core line probe assays and molecular analysis of untypeable samples."; Virus Res. 38:137-157(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                      Gaps
                                             K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update).
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus type lb.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
and molecular analysis of untypeable samples.";
Virus Res. 38:137-157(1995).
-- SDBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; L38348; AAC42168.1; -- InterPro; IPR002522; HCV_capsid.
PFam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 138; DB 12; Length 74;
Pred. No. 5e-13;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                          Score 138; DB 12; Length 74; Pred. No. 5e-13;
                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER 74 74 74 SEQUENCE 74 AA; 8412 MW; 2AE2DB0F976D10F3 CRC64;
                                                                                                                                                                                                                       74 AA; 8440 MW; 2AE2DB0F95BD10F3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                74 A.
                                                                                                                                                                                                                                                                                                 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Search completed: August 7, 2003, 11:19:02 Job time : 25.6364 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                          92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity 89.55,
Conservative
                                                                                                                                                                                                                                                                                                    25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             I_TaxID-31647;
                                                                                                                                                                                                 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
                                                                                                                                                                                  Polyprotein.
NON_TER
                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                   Best Local
```

us-09-491-146a-31.rai

```
ANTI-SENSE: NO
                                                                                                                                                                                                                                JS-08-921-887-31
Sequence 3, Appli
Sequence 10, Appl
Sequence 132, App
Sequence 14, Appl
Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 9, Appli
Sequence 2, Appli
Sequence 6, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Manager (1)

Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
                                                                         (without alignments)
112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 36, 2
                                                                August 7, 2003, 11:07:41; Search time 10:5455 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                            Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2
Sequence 9
Sequence 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 1
Sequence 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                             Issued_Patents_AA:*
    /cgn2_6/ptodata/1/laa/5A_COMB.pep:*
    /cgn2_6/ptodata/1/laa/5B_COMB.pep:*
    /cgn2_6/ptodata/1/laa/6A_COMB.pep:*
    /cgn2_6/ptodata/1/laa/6B_COMB.pep:*
    /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
    /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
    /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
       GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-921-887-31
US-08-290-658A-204
US-08-380-160-6
US-09-202-846-36
US-09-389-756-1
US-08-380-160-1
US-08-380-160-1
US-08-380-160-1
US-08-380-160-1
US-08-380-160-1
US-08-510-510-3
US-08-510-510-3
US-08-510-510-3
US-08-510-510-3
US-08-615-886C-198
US-08-615-886C-198
US-08-615-886C-198
US-08-615-886C-198
US-08-615-886C-198
US-08-615-886C-198
US-08-615-886C-198
US-08-615-886C-198
US-08-615-886C-198
US-08-616-886C-198
US-08-616-886C-198
US-08-616-886C-198
US-08-616-886C-198
                                                                                                                                                                                         al number of hits satisfying chosen parameters:
                                                                                                               150
1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                       328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                              protein search, using sw model
                                                                                                                                           BLOSUM62
Gapop 10.0 , Gapext
                                                                                                                                                                                                             length: 0
length: 2000000000
                                                                                                     US-09-491-146A-31
                                                                                                                                                                                                                                                                                                                                                                                                               Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sed
                                                                                                                Perfect score:
                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                                                                                                                            Minimum DB
Maximum DB
                                             OM protein
                                                                                                                         Sequence:
                                                                                                                                                                        Searched:
                                                                                                                                                                                                                                                                               Database
                                                                 Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
```

```
APPLICANT: KHUDYAKOV, YURI E.
TITLE OF INVENTION: MOSALC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAM
NUMBER OF SEQUENCES: 55
CORRESONED: ADDRESSE:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
STATE: GA
COUNTY: ALBANE
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
STATE: GA
COUNTY: ALSO CONTY: ALSO
COMPUTER: READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: INW PC COMPATIBLE
COMPUTER: THE PC COMPATIBLE
COMPUTER: PATENTIAN SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTION DATA:
SOFTWARE: PATENTIAN NUMBER: US/O8/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                          Арр
Арр
Арр
                                                                      Sequence 155, R
Sequence 156, R
Sequence 157, R
Sequence 158, R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                        Sequence 1
Sequence 1
Sequence 1
                    Sequence Sequence 1
                                                                                                                                            Sequence
Sequence
Sequence
Sequence
Sequence
Sequence
Sequence
                                                      Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 100.0%; Score 150; DB 3; Length 28; Best Local Similarity 100.0%; Pred. No. 5e-15; Matches 28; Conservative 0; Mismatches 0; Indels
                                               US-08-078-2718-1
US-08-290-665A-155
US-08-290-665A-155
US-08-290-665A-153
US-08-290-665A-159
US-08-290-665A-161
US-08-290-665A-161
US-08-290-665A-163
US-08-290-665A-163
US-08-290-665A-163
US-08-290-665A-163
                                                                                                                                                                                                                                                                                         US-08-290-665A-168
US-08-290-665A-169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
TELEPHONE: 404-818-3700
TELEPHONE: 404-818-3799
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 31, Application US/08921887 Patent No. 6030771 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Hepatitis virus US-08-921-887-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FRAGMENT TYPE: internal ORIGINAL SOURCE:
MOLECULE TYPE: protein HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
CLASSIFICATION: 435
```

```
Sequence 6, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
APPLICANT: DALBON, Pascal
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: FOR DETECTING THE LATTER
  SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 89.3%; Pred. No. 1.5e-12; Matches 25; Conservative 2; Mismatchee
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
                                                                                                                                                                                                  COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                      SSEE: MORGAN & FINNEGAN 1: 345 PARK AVENUE NEW YORK NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE: OLIFF & BERRIDGE STREET: P.O. Box 19928 CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE/DOCKET NUMBER: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEX: 421792,
INFORMATION FOR SEQ ID NO: 20,
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: homosapiens
INDIVIDUAL ISOLATE: SA6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SS: unknown unknown
                                        CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & F
                                                                                                                                                 ZUDE: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
TITLE OF INVENTION:
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: USA
ZIP: 22320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT-US95-10398-204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM:
                                                                                                   CITY: NEW
STATE: 'NE
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                           NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEPATITIS C VIRUS AND THE USE OF REAGENTS DERIVED FROM THESE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 139; DB 2; Length 191;
Pred. No. 1.5e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          R.H.
NUCLEOTIDE AND DEDUCED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY: ....
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FTI.NG DATE: 15-AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 204, Application PC/TUS9510398
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: BUKH, J., MILLER, R.H. AND
TITLE OF INVENTION: NUCLEOTIDE AND DEDUC:
TITLE OF INVENTION: AMINO ACID SEQUENCES:
TITLE OF INVENTION: CORE GENES OF ISOLATION: AND THE USE OF REAGE:
                     PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                            1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
                                                                                                                                                            Sequence 204, Application US/08290665A
Patent No. 5882852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 2026
REFERENCE/DOCKET NUMBER: 2026
TELECOMMUNICATION INFORMATION:
TELEFHONE: (212) 756-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 204:
                                                                                                                                                                                                                                                                                                                                                                                                     SSEE: MORGAN & FINNEGAN 345 PARK AVENUE NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match 92.7%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SA6
US-08-290-665A-204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRANDEDNESS: unknown
                                                                                                                                                                                                                                                         TITLE OF INVENTION: AN TITLE OF INVENTION: AN TITLE OF INVENTION: CO TITLE OF INVENTION: AN TITLE OF INVENTION: SEQUENCES: 2
                                                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NEW YORK
: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PCT-US95-10398-204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 3
                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                     င်
```

Gaps

```
APPLICANT: DALBON, Pascal
APPLICANT: JOLIVET, MIChal
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALL)
TITLE OF INVENTION: FOR DETECTING THE LATTER
CORRESPONDENCES: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                            Score 138; DB 4; Length 43;
Pred. No. 4.1e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                 Query Match 92.0%; Score 138; DE Best Local Similarity 89.3%; Pred. No. 4.1¢ Matches 25; Conservative 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
CLASSIFICATION: 530
PROOF APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER REALABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
FILING DATE: 09-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Wegner, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 0534(
TELECOMMUNICATION INFORMATION:
TELEPAX: (202) 672-5300
TELEPAX: (202) 672-5300
TELEPAX: (202) 672-5300
TELEPAX: 402 anino acids
LENGTH: 43 anino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-380-160-2; Sequence 2, Application US/08380160; Patent No. 6235284; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: OLIFF & BERRIDGE STREET: P.O. BOX 19928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE/DOCKET NUMBER: WP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 44 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEX:
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MOLECULE TYPE: peptide HYPOTHETICAL: NO ANTI-cent
                                                                                                                                                                                                                                                                                                                  ; MOLECULE TYPE: peptide US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 USA
                                                                                                                                                                                                                                                                          STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22320
                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER: Floppy disk
COMPUTER: IBM PC COMPATION STREET
COMPUTER: IBM PC COMPATION STREET
SOFFWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
           OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.0%; Score 138; DB 3;
illarity 89.3%; Pred. No. 3.2e-13;
Conservative 2; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
APPLICANT: YAMAGUCHI, Kenjiro
APPLICANT: KASHIWAKUMA, TOMIKO
APPLICANT: CHIBA, YUKie
APPLICANT: YAGI, Shintaro
APPLICANT: YAGI, Shintaro
APPLICANT: HASEGAMA, AKITA
TILLE OF INVENTION: CHIMERA ANTIGEN PEPTIDE
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                               CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: BEATIGHS, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: 30,024
TELECHONE: (703)836-6400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: Human Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: US/09/020,846
FILING DATE: 09-FEB-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 9-027015
FILING DATE: 10-FEB-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA;
APPLICATION NUMBER: JP 8-024045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 36, Application US/09020846
Patent No. 6322965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-terminal
                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acids
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 25; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE: 09
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ANTI-SENSE: NO
FRAGMENT TYPE: N-
ORIGINAL SOURCE:
                                                                                                    FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

```
RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: JOLIVET, MICHEL
APPLICANT: JOLIVET, MICHEL
APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
TITLE OF INVENTION: TREATING AN HCV INFECTION
TITLE OF INVENTION: TREATING AN HCV INFECTION
CURRENT APPLICANTION NUMBER: US/09/389,756
CURRENT FILING DATE: 1999-09-07
EARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PALCHLIN Ver. 2.1
                                                                                                                                                                                                                                                                        ö
                                                                                                                                 /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    the H Strain
                                                                                                                                                                                                                              Score 138; DB 3; Length 44;
Pred. No. 4.2e-13;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 138; DB 4;
Pred. No. 4.2e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PUBLICATION INFORMATION:
AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mismatches
                                                                                                                                                                                                                                                                                                        1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                            1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORIGINAL SOURCE:
ORGANISM: Human Hepatitis C Virus
STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                vatent No. 6576240 TTTLE: of Hepatitis Virus JUNRNAL: Proc. Natl. Acad. Sci. U.S.A. VOLUME: 88
                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 1, Application US/09389756
Patent No. 6576240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 1, Application US/08380160 Patent No. 6235284 GENERAL INFORMATION:
                                                                                                                                                                                                                                92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: DALBON, Pascal APPLICANT: JOLIVET, Michel
FRAGMENT TYPE: N-terminal
                                                                                                                                                                                                                 Ouery Match
Best Local Similarity 89.39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RELEVANT RESIDUES: 2 TO 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 92.0
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                        NAME/KEY: Peptide
LOCATION: 1.44
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PAGES: 3392-3396
                                                                                                                                                                                                                                                                                                                                                                                                       JLT 7
09-389-756-1
                                                                                                                                                                                         US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 8
US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 44
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQ ID NO 1
                                                                          FEATURE
                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
NS-07-946-054-9
Sequence 9, Application US/07946054
Sequence 9, Application US/07946054
Sequence 9, Application US/07946054
Sequence 9, Application US/07946054
Sequence 9, Application SS8234
SEMERAL INFORMATION:
APPLICANT: Wang, Chang Y1
APPLICANT: Wang, Chang Y1
TITLE OF INVENTION: No. 5582968el Branched Hybrid and Cluster
TITLE OF INVENTION: No. 5582968-B Hepatitis
TITLE OF INVENTION: No. 5582968-B Hepatitis
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
  SYNTHETIC POLYPEPTIDES BELONGING TO THE HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY FOR DETECTING THE LATTER 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                              COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.0%; Score 138; DB 3;
ilarity 89.3%; Pred. No. 4.3e-13;
Conservative 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNINRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERNCE/DOCKET NUMBER: WPB 28682
TELECOMMUNICATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-YAX-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORIGINAL SOURCE:
ORGANISM: Human Hepatitis C Virus
TITLE OF INVENTION: SYNTHETIC PORTITIES OF INVENTION: HERATITIS CALLE OF INVENTION: FOR DETECTION WIMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
STREET: P.O. BOX 19928
                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FRAGMENT TYPE: N-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: Peptide
LOCATION: 1..45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OTHER INFORMATION:
                                                                                                                                                              Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                            USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
                                                                                                                                                                                                                               22320
                                                                                                                                                                                         Ϋ́
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ANTI-SENSE:
                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-380-160-1
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 61;
                                                                                                                                                                                                                                                                                                                      DB 1; Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 3, Application US/08530550

Patent No. 573631

GENERAL INFORMATION: Barbara
APPLICANT: Wang, Chang Yi

TITLE OF INVENTION: Peptides Effective for Diagnorman Properties of Type of Ty
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 92.0%; Score 138; DB 1.
Best Local Similarity 89.3%; Pred. No. 6e-13;
Matches 25; Conservative 2; Mismatches
                                                                                                                                                                                                                                                                                                                   Score 138; DB 1;
Pred. No. 6e-13;
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; Sequence 26, Application US/08262037; Patent No. 5747239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               34,045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE/DOCKET NUMBER: 20
TELECOMMUNICATION INFORMATION:
                             TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO: 23:
                                                                                                                                                                                                                                                                                                                Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATTORNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEFAX: (516)273-1717 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS:
                                                                                            SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; MOLECULE TYPE: peptide US-08-530-550-3
                                                                                                                                                                            TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION: 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hauppauge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 23, Application US/08083947

Patent No. 5639594

GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
APPLICANT: Hosein, Barbara
TITLE OF INVENTION: No. 5639594el Linear and Branched Peptides Effective
TITLE OF INVENTION: in Diagnosing and Detecting No. 5639594-B Hepat
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1; Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
COMPUTER: IBM PC Compatible
COMPUTER: IBM PC COMPATIBLE
CONFING SYSTEM: PC-DOS/MS-DOS
SOFFWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,947
FILING DATE: 19930628
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 946,054
FILING DATE: 15-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
NAME: WALLSON, M. Lisa
                                                                                                                                                                                                                                                                                                   OURTHAING SYSTEM: PC-DUS/MS-LUSS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/O7/946,054
FLING DATE: 15-SEP-1992
CLASSIFICATION: 435
ATTONNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/DOCKET NUMBER: 2000
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION OF 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 aming acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 138; DB 1,
Pred. No. 6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
ADDRESSEE: United Biomedical Inc.
STREET: 25 Davids Dr.
CITY: Hauppauge
STATE: New York
                                                                                                                                                              ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE/DOCKET NUMBER: 2000Z
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REGISTRATION NUMBER: 34045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 92.0
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; MOLECULE TYPE: peptide US-07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hauppauge
                                                                                                                                 USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: US
ZIP: 11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 10
US-08-083-947-23
                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: TBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 33. Application PC/TUS9407088
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
APPLICANT: Woein, Barbara
TITLE OF INVENTION: Novel Linear And Branched
TITLE OF INVENTION: Diagnosing And Detecting
TITLE OF INVENTION: Non-A, Non-B Hepatitis
NUMBER OF SEQUENCES: 23
CORRESSEDE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: New York
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          92.0%; Score 138; DB 5; 89.3%; Pred. No. 6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: PCT/US94/07088
FILING DATE: 22-JUNE-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPRECT 5.1
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/083,947
FILING DATE: 28-JUNE-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                           TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-273-2828
TELEFAX: 516-273-1717
                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: M. Lisa Wilson
REGISTRATION NUMBER: 34,6
REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: LIN, MARIA C.H. REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CLASSIFICATION:
                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PCT-US94-07088-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PCT-US93-08638-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                           SYNTHETIC PEPTIDES SPECIFIC FOR
THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
INFECTION AND PREVENTION THEREOF AS VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 9, Application PC/TUS9308638
GENERAL INFORMATION:
APPLICANT: United Biomedical Inc.
TITLE OF INVENTION: Novel Branched Hybrid and Cluster Peptides
TITLE OF INVENTION: Effective in Diagnosing and Detecting Non-A,
TITLE OF INVENTION: Non-B Hepatitis
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
CLASSIFCATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/719,819
FILING DATE: 24-June-1991
APPLICATION NUMBER: 07/667,275
FILING DATE: 11-Mar-1991
APPLICATION NUMBER: 07/651,735
FILING DATE: 07-Eb-1991
APPLICATION NUMBER: 07/558,799
FILING DATE: 26-July-1990
APPLICATION NUMBER: 07/510,153
FILING DATE: 16-April-1990
APPLICATION NUMBER: 29,323
FILING DATE: 16-April-1990
APPLICATION NUMBER: 29,323
FILING DATE: 21-758-4800
TELEPHONE: 212-758-4800
TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 11-1-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB 1
Pred. No. 6e-13;
                   APPLICANT: Chang Y1 Wang and Barbara Hosein TITLE OF INVENTION: SYNTHEIT CPEPTIDES SPECITIONS TITLE OF INVENTION: THE DETECTION OF ANTIBOUTILE OF INVENTION: INFECTION AND PREVENTION NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: WORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.0%; Scur.
89.3%; Pred. No. vc.
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 PKPORKTKRNINRRPODVKFPGGGOIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                    OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: WORDPERFECT 5.1 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/262,037
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE: UNITED BIOMEDICAL INC. STREET: 25 Davids Drive CITY: Hauppauge STATE: New York
                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     iery Match 52.0.
est Local Similarity 89.30
Matches 25; Conservative
                                                                                                                                                                              345 PARK AVE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unknown
                                                                                                                                                                                                                                                                                                                                                               OPERATING SYSTEM:
GENERAL INFORMATION:
                                                                                                                                                                                                                              NEW YORK
                                                                                                                                                                                                         CITY: NEW YORK
                                                                                                                                                                                                                                                    COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; TOPOLOGY:
US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
PCT-US93-08638-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

INFORMATION FOR SEQ ID NO:

ö

Gaps

us-09-491-146a-31.rai

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: Hosein, Barbara
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Peptides Effective for
TITLE OF INVENTION: Diagnosis and Detection of Hepatitis C Infection
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0; Gaps
                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                              .
0
                                                                                                                                                               Query Match 92.0%; Score 138; DB 5; Length 61; Best Local Similarity 89.3%; Pred. No. 6e-13; Matches 25; Conservative 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92.0%; Score 138; DB 5; Length 61; 89.3%; Pred. No. 6e-13; 1: Indels 1ive 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE: Morgan & Finnegan, L.L.P
STREET: 345 Park Avenue
CITY: New York
STRATE: New York
COUNTRY: USA
ZIP: 10154
COMPUTER: TIMP CCOMPATIBLE
COMPUTER: TIMP PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13660
FILING DATE: 23 October 1995
CLASSIFCATION:
APPLICATION NUMBER: 08/333,573
FILING DATE: 01 November 1994
ATTORNEY AGENT INFORMATION:
NAME: MATIA C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4118PC
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                         1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                              4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 3, Application PC/TUS9513660 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acids
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 89.3*
Matches 25; Conservative
                                                           ; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US94-07088-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                    RESULT 15
PCT-US95-13660-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PCT-US95-13660-3
                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                           ò
```

Search completed: August 7, 2003, 11:23:53 Job time : 11.6364 secs

```
APPLICANT: DOLIVET, MICHEL
APPLICANT: DOLIVET, MICHEL
APPLICANT: DOLIVET, MICHEL
APPLICANT: DOLIVET, MICHEL
APPLICANT: DADAVIERE, LAURENT
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILE REFERENCE: 103959
CURRENT APPLICATION NUMBER: US/10/367,677
CURRENT FILING DATE: 1999-09-07.
PRIOR PILING DATE: 1999-09-07.
PRIOR PELING DATE: 1999-09-07.
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/FR98/00442
PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 1
LENGTH: 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PUBLICATION INFORMATION:
AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
TITLE: Of Hepatitis Virus
US-09-929-955-2
US-09-851-138-76
US-09-851-138-76
US-09-851-138-12
US-09-851-138-12
US-09-895-046-55
US-09-899-046-54
US-09-899-046-54
US-09-899-046-54
US-09-878-281-54
US-09-878-281-54
US-09-973-025-50
US-09-973-025-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-952-572-9
US-09-742-659-4
US-09-742-659-4
US-09-742-659-4
US-09-742-659-4
US-09-742-659-4
US-09-952-572-9
US-09-952-572-9
US-09-952-572-9
US-09-952-572-9
US-09-952-972-9
US-09-952-975-10
US-09-952-975-10
US-09-952-975-10
US-09-952-975-10
US-09-952-975-10
US-09-952-975-10
US-09-952-975-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL: Proc. Natl. Acad. Sci. U.S.A. VOLUME: 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/10367677 Publication No. US20030118604A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RELEVANT RESIDUES: 2 TO 45 US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                            2985
3011
3011
3011
3011
3011
3012
        PAGES: 3392-3396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-10-367-677-1
        TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
          Sequence 1, Appli
Sequence 7, Appli
Sequence 17, Appl
Sequence 14, Appl
Sequence 78, Appl
Sequence 60, Appl
Sequence 61, Appl
Sequence 152, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Appli
Sequence 10, Appl
                                                                                                                                                              (without alignments)
231.506 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                             August 7, 2003, 11:20:11; Search time 14.3636 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Published_Applications_Ash.

/ cgn2_6/ptodata1/pubpaa/US07_PUBCCMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US07_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US06_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US06_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US07_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US08_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US08_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US08_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US08_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US09_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US10A_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US10A_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US10_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US10_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US10_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US10_NEW_PUB.pep:*
                         5.1.6
Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-756-875-8
US-09-921-397-77
US-09-921-397-78
US-09-851-138-46
US-09-851-138-60
US-09-871-138-60
US-09-878-281-152
US-09-899-046-42
US-09-899-046-44
US-09-878-281-44
                                                                                                                                                                                                                                                                                                                                                                                                        al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                150
1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                     451899 segs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                         GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                          BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                             US-09-491-146A-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0000000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               44
74
91
97
103
108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         137
138
166
169
169
169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                    score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  138
1138
1138
1138
1138
1138
1138
1138
                                                                                                          protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Database :
                                                                                                                                                                                                                                                                  Sequence:
                                                                                                                                                                                                                                                                                                                                                                     Searched:
                                                                                                                                                                                                                             Title:
Perfect :
                                                                                                                                                Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ę
```

Sequence 4, Appli Sequence 9, Appli Sequence 1, Appli Sequence 20, Appl

Sequence 3, Appli Sequence 1, Appli Sequence 20, Appl

Sequence Sequence

Length 44;

Appl Appl Appl Appl Appl

Sequence 54, Sequence 54, Sequence 50, Sequence 50, Sequence 50, Sequence 50, Sequence 23, Sequence 23, Sequence 40,

Appl Appl Appl

Sequence 12, Sequence 1 Sequence 5 Sequence 5

Sequence

Sednence

Appl

Sequence Sequence Sequence

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 97;
                                                                                                                                                                                                                  Length 91;
                                                                                                                                                                                                                                                           1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 10;
                                                                                                                                                                                                                  Score 138; DB 9;
Pred. No. 2.8e-12;
2; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 8, Application US/09756875
Patent No. US20020150990A1
GENERAL INFORMATION:
APPLICANT: PIKE, IAN
TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB 10
Pred. No. 3e-12;
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE: Suite 701-E Columbia Square STREET: 555 13th Street, N. W.
                                                                                                                                                                                                                                                                                             1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                    5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPORKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-406-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE/DOCKET NUMBER: 1808-157A TELECOMMUNICATION INFORMATION: TELEPHONE: (202)783-6040
                        60/092,339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,875
CURRENT FILING DATE: 2001-01-10; PRIOR APPLICATION NUMBER: 60/092; PRIOR FILING DATE: 1999-07-10; NUMBER OF SEQ ID NOS: 5
SOFWARE: Patentin version 3.0; SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                92.0%;
hilarity 89.3%;
Conservative 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92.0%;
89.3%;
                                                                                                                                                       ) ORGANISM: Hepatitis C Virus US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 89.39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEC ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                               Query Match
Best Local Similarity
..... 25; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STATE: D. C. COUNTRY: U. S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20004
                                                                                                                                                                                                                                                                                                                                                                                                              US-09-756-875-8
                                                                                                                                          TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CITY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 14114.0349U2
CURRENT APPLICATION NUMBER: US/09/758,308
                                                                                                                                                                                                                                                                       STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                        ö
                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRETING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRET APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKnown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 74;
                      Indels
Similarity 89.3%; Pred. No. 1.3e-12; 25; Conservative 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.0%; Score 138; DB 10;
89.3%; Pred. No. 2.2e-12;
tive 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                   CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE STREET! P.O. BOX 4433
CITY: HOUGYON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: KAMMERRY, PATRICIA A,
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-138-10
                                                           1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                               1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                           Sequence 10, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 1, Application US/09758308 Patent No. US20020090607A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 74 amino acids
                                                                                                                                                                                                                                                     APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                  AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 92.0°
Best Local Similarity 89.3°
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO
                                                                                                                                                                           US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-758-308-1
  Best Local
Matches 2
                                                                                                                                                           RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 3
                                                           ð
                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

ő

```
APPLICANT: MAERTENS, GEERT STUVER, LIEVEN STUVVER, LIEVEN TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                           polypeptides selected from
the hepatitis C virus and
                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 113;
                                                                                     Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Microsoft Word 6.0 / ASCII text output CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB 10;
Pred. No. 3.5e-12;
2; Mismatches 1;
                                                                                 Score 138; DB 10;
Pred. No. 3.3e-12;
2; Mismatches 1;
                                                                                                                                                                                                                                                                                           US-09-91-1997-78
Sequence 78, Application US/09921397
Fatent No. US20020151484A1
GENERAL INFORMATION:
TITLE OF INVENTION: SID nucleic acids and pol TITLE OF INVENTION: pathogenic strain of the TITLE OF INVENTION: applications thereof FILE REFERENCE: B4809A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 78
LENGTH: 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                1 PKPORKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 46, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AGENTS
                                                                                 Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; ORGANISM: Hepatitis C virus US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: USA
ZIP: 77210-4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OF HEPATITIS C VIRUS GENOTYPES
AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                 TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a TITLE OF INVENTION: Dathogenic strain of the hepatitis C virus and TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and TITLE OF INVENTION: applications thereof FILE REFERENCE: B4809A - Jax CURRENT APPLICATION NUMBER: US/09/921,397 CURRENT APPLICATION NUMBER: US/09/921,397 PRIOR APPLICATION NUMBER: E004002225,7 PRIOR APPLICATION NUMBER: E004002225,7 PRIOR FILING DATE: 2000-08-03 NUMBER OF SEQ ID NOS: 156 SOFTWARE: Patentin Ver. 2.1 SEQ ID NO 77 LENGTH: 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 138; DB 10;
Pred. No. 3.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: 08/836,075
FILING DATE: «CORNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTRNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES
AND THEIR USE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 14, Application US/0951138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
                                  Sequence 77, Application US/09921397
Patent No. US20020151484A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES: 20 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                          TYPE: PRT
ORGANISM: Hepatitis C virus
09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE: TEXAS
                                                                            GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 6
US-09-851-138-14
             US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92.0%; Score 138; DB 11; Length 166; 89.3%; Pred. No. 5.3e-12; 1ve 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 11; Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DAPA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM PC_COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
            Mismatches
                                                                   1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                  1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92.0%; Score 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US/09/878,281
                                                                                                                                                                                                                                                                        ; Sequence 152, Application US/09899046; Publication No. US20030008274A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              s: Floppy disk
IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; Sequence 152, Application US/09878281; Publication No. US20030032005A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
         25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; MOLECULE TYPE: protein US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                          TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
US-09-878-281-152
                                                                                                                                                                                                                                               US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
            Matches
                                                                   δ
                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: MAERTENS, GEERT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: NAT THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CORRESPONDENCE ADDRESS: CORRESPONDENCE ADDRESS: CORRESPONDENCE ADDRESS: CORRESPONDENCE ADDRESS: CONTY: HOUSTON CITY: HOUSTON CITY: HOUSTON CITY: HOUSTON CITY: HOUSTON COMPUTE: TEXAS COUNTRY: USA COMPUTE: TEMPS COMPUTE: US/09/851,138 FILING DATE: O9-MSY-2001 PRIOR APPLICATION NUMBER: EP 94870166.9 FILING DATE: CUNROWN CONTROL OF ATTORNATION COMPUTE: EP 95870076.7 FILING DATE: 28 Jun 1995 ATTORNEY AMANGE TO THE ADDRESS COMPUTE: TEMPS COMPU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 138; DB 10; Length 137;
Pred. No. 4.3e-12;
2; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 10;
Pred. No. 4.3e-12;
FILING DATE: 21 OCT 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
RECISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: 1NNS:004
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
NO FOR SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear;
HOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-851-138-60; Sequence 60, Application US/09851138; Publication No. US20020183508A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                           LENGTH: 137 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 138 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92.0%;
ilarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 25; Conserval
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR
```

Gaps

ô

```
New sequences of hepatitis C virus
genotypes for diagnosis, prophylaxis and therapy.
270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New sequences of hepatitis C virus
genotypes for diagnosis, prophylaxis and therapy.
270
                   Score 138; DB 11; Length 169; Pred. No. 5.4e-12;
                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Indels
                                                                                                                                                                                                                                                                                                                                  COMPUTER: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: PC.005/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
APPLICATION DATA:
FILING DATE:
PRIOR BOTE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPRATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 138; DB 11;
Pred. No. 5.4e-12;
2; Mismatches 1;
                                                        2; Mismatches
                                                                                              1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                               5 PKPORKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 44, Application US/09878281 Publication No. US20030032005A1 GENERAL INFORMATION:
                                                                                                                                                                                                                       Sequence 42, Application US/09878281; Publication No. US20030032005A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.0%;
                   92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               : 169 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          169 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                        Conservative
                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: New TITLE OF INVENTION: GENCINDER OF SEQUENCES: 270 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy di
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE: PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
TITLE OF INVENTION: No TITLE OF INVENTION: 90 NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                 Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; MOLECULE TYPE:
US-09-878-281-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                          US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-878-281-44
                                                                                                                                                                                                                                                                                    APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGIH:
                                                                                                                                                                                      RESULT 14
                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/NS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                  of hepatitis C virus diagnosis, prophylaxis and therapy.
                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 169;
                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                            TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylax NUMBER OF SEQUENCES: 270
COMPUTER PARABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
. 5.3e-12;
1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 138; DB 11;
Pred. No. 5.4e-12;
2; Mismatches 1;
                 Pred. No. 5.36
; Mismatches
                                                                                          PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPORKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                    ; Sequence 42, Application US/09899046; Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-899-046-44
Sequence 44, Application US/08899046
'ublication No. US20030008274A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                   7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO: 44: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
               88.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENCTH: 169 amino acids
TYPE: amino acid
TOPOLOGY: linear
                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: protein US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
               Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                                        US-09-899-046-42
                                                                                                                                                                  RESULT 12
                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ó
```

ö

Gaps

```
ő
                                                        0; Gaps
Query Match 92.0%; Score 138; DB 11; Length 169; Best Local Similarity 89.3%; Pred. No. 5.4e-12; Matches 25; Conservative 2; Mismatches 1; Indels C
```

ò

q

Search completed: August 7, 2003, 12:01:14 Job time : 14.3636 secs

Scoring table:

Searched:

Minimum DB Maximum DB

Database

Title: Perfect score:

Sequence:

OM protein

Run on:

```
Encoded by Hepatit
Encoded by Hepatit
NC mosalc protein
Blood transmiscibl
HCV-SI full-length
                                                                                                                                                                Non-A, Non-B Hepat
Non-A, Non-B Hepat
Hepatitis C virus
Sequence of peptid
Fragment of open r
Hepatitis C virus
HCV capsid beptide
HCV capsid peptide
Hepatitis C virus
Hepatitis C virus
                                                                                Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                       Human hepatitis C
Human hepatitis C
HCV core-envelope
HCV core-envelope
HCV core-envelope
HCV core-envelope
HCV core-envelope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Non-A, Non-B Hepat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Peptide VIIIE base
Anti-HCV antibody
Prototype peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mosaic protein, antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NC mosaic protein amino acid fragment H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                                                                                                                                                                                                                                                                                                                                       AAR54065
AAW37380
                                                                                                                                                                                                                                                                                                                                                                                                        AAY26952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR69545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY06680 standard; Protein; 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Fields HA, Khudyakov YE;
 470
470
3010
3010
3010
40
40
50
82
82
191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1999-204671/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               W09910506-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              17-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              04-MAR-1999.
 AAY06680;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 1
AAY06680
 NC mosaic protein
NC mosaic protein
HCV type 3 capsid
Hepatitis C virus
Encoded by Hepatit
                                                                                                               Search time 38.5455 Seconds (without alignments)
115.301 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | SIDSI/ggdata/geneseq/geneseqp-embl/AA1981_DAT:
| SIDSI/ggdata/geneseq/geneseqp-embl/AA1981_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1985_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1985_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1985_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1990_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2000_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2001_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2002_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003_DAT:
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003_DAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       // rembl/AA1997.DAT:*
//gcgata/geneseqp-embl/AA1998.DAT:*
//gcgata/geneseq/geneseqp-embl/AA1999.DAT:*
//gcgata/geneseq/geneseqp-embl/AA2000.DAT:*
//gcgata/geneseq/geneseqn-embl/AA2000.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.
                                                                                                                                                                                                                                                                                                                                    1107863
               GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                    tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                   151
1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV 28
                                                                                                                                                                                                                                                                                                    1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                  August 7, 2003, 11:05:37;
                                                                                                                                                                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY06680
AAY06673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB71258
AAR92968
AAR92971
AAR92971
AAR92970
AAR92970
                                                                                                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                      seq length: 0
seq length: 200000000
                                                                                                                                                                                   US-09-491-146A-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   220
220
220
24
71
71
71
71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28
28
189
191
191
191
191
191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 88.7
88.7
88.7
88.7
88.7
88.7
```

Score

Result Š 459786

ő

Gaps ö

;-;

Length 28; Indels

```
This invention describes a novel attenuated flavivirus live vaccine comprising a flavivius mutant that has a deletion of at least 4 consecutive amino acids from the capsid protein, provided that the capscutive amino acids from the capsid protein, provided that the creminal hydrophobic region is not affected by the deletion. The vaccine of the invention has virucide, antinflammatory and hepatotropic activity. The attenuated vaccine, and similar nucleic acid vaccines that encode the mutated capsid protein, are useful for protection against a vide range of flavivirus diseases, e.g. yellow fever, Japanese encephalitis, dengue, classical swine fever, bovine viral diarrhoea and hepatitis C. The specified deletion: (i) produces a reliably attenuated virus that does not revert to virulence; (ii) is exactly defined and does not effect immune responses to important proteins; and (iii) can not concert from contrarial virus by recombination. The mutant viruses eliminate the need to produce large amounts of infectious/virulent viruses, and can be produced with less expense. The protective response to flavivirus lasts significantly longer than that to killed vaccines. This sequence represents a fragment of the capsid protein from Heptitis c virus (HCV) type I described in the disclosure of the invention.
artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences AAY06073-681 represent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Attenuated flavivirus live vaccine, useful for protection against e.g. yellow fever, comprises virus with attenuating deletion of amino acids from the capsid protein -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Capsid protein; attenuated vaccine; virucide; antinflammatory; hepatotropic; yellow fever; Japanese encephalitis; dengue; classical swine fever; bovine viral diarrhoea; hepatitis C.
                                                                                                                                                     Score 134; DB 20;
Pred. No. 2.7e-12;
0; Mismatches i;
                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV type 3 capsid protein fragment.
                                                                                                                                                                                                                                                                                                                                                                AAB71258 standard; protein; 189 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; Fig 2; 30pp; German.
                                                                                                                                                       88.7%;
96.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11-FEB-2002; 2002WO-AT00046.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21-FEB-2001; 2001AT-0000272.
                                                                                                                                 Ouery Match
Best Local Similarity 90.2.
Best Local Similarity 90.2.
                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2002-667064/71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mandl C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (HEIN/) HEINZ F X. (MAND/) MANDL C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis c virus.
                                                                                                                   28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200266621-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                              18-NOV-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         29-AUG-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Heinz FX,
                                                                                                                                                                                                                                                                                                                                                                                                        AAB71258;
                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                RESULT 3
 8$$$$$$$$
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                 The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted rigation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosalc protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosalc protein and the artificial mosalc protein are useful for detecting a hepatitis infection in an individual. The mosalc gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The invention relates to a mosaic protein, comprising a plurality of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                  Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                              Score 151; DB 20;
Pred. No. 9.2e-15;
Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NC mosaic protein amino acid fragment A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (USSH ) US DEPT HEALTH & HUMAN SERVICES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ź
                                                                                                                                                                                                                                                                                                       comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY06673 standard; Protein; 28
                                                                                                                                                                                                                                                                                                                                                                                100.0%;
                 Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 5; Fig 9; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97US-0921887.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17-JUN-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                uery Match 100.

St Local Similarity 100.

atches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1999-204671/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                            28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO9910506-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   04-MAR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fields HA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY06673;
                                                                                                                                                                                                                                                                                                                                                Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
```

Matches

셤

S

AAR92968
IID AAR9
XX
AC AAR9
XX
XX
AC AAR9
XX
YX
YY
Hepa

```
envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                             core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated CDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the CDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, Iymphocytes or other monounclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNA and amino acid sequence of HCV envelope 1 and core proteins used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 134; DB 17; Length 191;
Pred. No. 2e-11;
0; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus isolate DK12 core protein.
                                                     Hepatitis C virus isolate S52 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                US DEPT HEALTH & HUMAN SERVICES, US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PKPORKTKRNTIRRPODVKFPGGGOI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR92971 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 4; Page 208; 340pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 88.7%;
Best Local Similarity 96.2%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95WO-US10398.
                                                                                                                                                                                                                                                                                                                                                       95WO-US10398,
                                                                                                                                                                                                                                                                                                                                                                                                          94US-0290665
32-OCT-1996 (first entry)
                                                                                                           ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1996-139709/14
                                                                                                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 191 AA;
                                                                                                             envelope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAT16643
                                                                                                                                                                                                                                                                                                                                                                                                       15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO9605315-A2
                                                                                                                                                                                                                                               W09605315-A2
                                                                                                                                                                                                                                                                                                                                                       15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15-AUG-1994;
                                                                                                                                                                                                                                                                                                  22-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 02-0CT-1996
                                                                                                                                          hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR92971;
                                                                                                           E1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV; El;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR92971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NO COCCOCC COCCOCC COCCC COCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other monounclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection
                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 134; DB 17; Length 191; Pred, No. 2e-11;
                                                        Length 189;
                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Indels
                                                     DB 23;
                                                  Score 134; DB 23
Pred. No. 2e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus isolate HK10 core protein.
                                                                                                        0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                       4 PKPQRKTKRNTIRRPQDVKFPGGGGI 29
                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGVI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR92969 standard; Protein; 191 AA.
                                                                                                                                                                                                                                                                                                                                                       Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 4; Page 207; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                  AAR92968 standard; Protein; 191
                                                  88.7%;
96.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Purcell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Match 88.7%;
Local Similarity 96.2%;
es 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       95WO-US10398
                                                                      Local Similarity 96.2 nes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1996-139709/14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              191 AA;
189 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAT16642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22-FEB-1996,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hepatitis.
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                       AAR92968;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV; E1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR92969
```

Matches

ò 셤

388888

RESULT 5
AAR92969
ID AAR9
XX
AC AAR9

ö

Gaps

4

```
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR92970;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR92970
ID AAR9
XX AC AAR9
XX DT 02-C
XX XX HPP
XX HCV
XX KW HCV
XX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note- "amino acid in this position is designated X in the specification, but codon usage shows that the only possible amino acid at this pos. is Gly"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /note= "amino acid in this position is designated X in
the specification, but codon usage shows that the
only possible amino acid at this pos. is Gly"
                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus; subtype; polymerase chain reaction; amplification; PCR; primer; probe; antibody; infection.
                                                                                                                                AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other monounclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                               DNA and amino acid sequence of HCV envelope 1 and core proteins -used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus types 7c(8a) isolates VN4 amino acids 1-317
                                                                                                                                                                                                                                                   DB 17; Length 191;
                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                   Score 134; DB 17
Pred. No. 2e-11;
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 144..149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 'label = Met, Thr, Lys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /label- val, Ala, Glu, Gly
171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /label- Cys, Arg, Ser, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      label- Val, Ala, Asp, Gly
(USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                             Claim 4; Page 209-210; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /label= Met, Leu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /label= Ser, Arg
Misc-difference 167
                              Miller RH, Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                AAR96547 standard; peptide; 319
                                                                                                                                                                                                                                                 88.7%;
96.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /label- Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /label- Gly
                                                                                                                                                                                                                                                                       Conservative
                                                 WPI; 1996-139709/14. N-PSDB; AAT16645.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 172
                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Key
Misc-difference 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 16]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 15
                                                                                                                                                                                                                              191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                        10-MAR-1997
                                                                                                                                                                                                                               Sequence
                             Bukh J,
                                                                                                                                                                                                                                                                                                                                                      ζ
FT
                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                             g
```

```
The sequences AAR96526-R96578 represent novel sequences isolated from hepatitis C virus subtypes different from subtypes la-c, 2a-d, 3a-f, 4a-j, 5a and 6a. They esp. from the novel subtypes la-f, 2e-i, 2k, 2l, 3q-jk.m. and 6a. They esp. from the novel subtypes la-f, 2e-i, 2k, 2l, 3q-jk.m. and 6a. They esp. from the novel subtypes la-f, 2e-i, 2k, 2l, 3q-jk.m. and 2a-jk.m. and 10 m. sequence corresp. to the 5' untranslated region (UR), the Core/El, NS4 or NS5B regions of the genome. This sequence represents amino acids l-317 from the HCV types 7c and 8a isolates VN4.

The new HCV types were isolated from patients with chronic HCV from the Benelux countries, France, Cameroon and Vietnam, because of their abeniux countries, France, Tameroon and Vietnam, because of their abenium the countries of the NNA was extracted, cDNA synthesized and PCR amplified, cloned and genotyped. The 5'UR, Core-El and NSSB regions were sequenced either directly or partially and used to classify the new viruses into (sub)types based on comparison with known sequences. The sequences can also be used to synthesize probes and primers for the sequences can also be used to synthesize probes and primers for the sequences can also be used to synthesize probes and primers for the catection of HCV in a sample. The polypeptides can be used to detect anti-HCV antibodies, for HCV typing or to prevent HCV infections.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sub:type
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus poly:nucleic acid unique to unidentified - used to develop probes and primers for new sub:types and to prevent and treat infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  88.7%; Score 134; DB 17; 96.2%; Pred. No. 3.5e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Mismatches
                                    Leu, Ile, Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus isolate S2 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                       Val
                                                                                                                                                                                                                                         /label- Asn, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR92970 standard; Protein; 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 25; Fig 3; 150pp; English
                                'label- Phe,
                                                                                                      /label- Phe,
                                                                                                                                                                    /label- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                      95WO-EP04155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       95EP-0870076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity 96.2
hes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (INNO-) INNOGENETICS NV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Maertens G, Stuyver L;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1996-251460/25
Misc-difference 174
                                                                 Misc-difference 177
                                                                                                                                                                                                        Misc-difference 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             319 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; AAT27958
                                                                                                                                   Misc-difference
                                                                                                                                                                                                                                                                                                           WO9613590-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                  23-OCT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-JUN-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-OCT-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 02-OCT-1996
```

S

ö

Gaps

ö

Indels

Length 470;

```
CDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-B. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV. See AAQ40425-040439.
                                                cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-A. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV. See AAQ40425-Q40439.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV; non-A, non-B hepatitis virus; NANBHV; liver disease;
polymerase chain reaction; diagnostic method.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DNA and cDNA of hepatitis C virus - useful as probes for diagnosing HCV infection
                                                                                                                                                                                        Score 130; DB 14;
Pred. No. 2e-10;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86.1%; Score 130; DB 14; 92.3%; Pred, No. 2e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. 2e-1
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Encoded by Hepatitis C virus clone JK3-B.
                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
              Claim 4; Page 26-28; 44pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 4; Page 28-30; 44pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                            Ŕ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (SANW ) SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                                                                                          AAR34474 standard; Protein; 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR34475 standard; Protein; 470
                                                                                                                                                                                            86.18;
92.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   91JP-0153736
                                                                                                                                                                           Query Match
Best Local Similarity 92.3v
Local 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                              10-JUL-1993 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best_Local Similarity 92.3
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1993-130638/16
                                                                                                                                                           470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAQ40432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JP05068562-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23-MAR-1993.
                                                                                                                                                                                                                                                                                                                                                                                                                               AAR34474;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                         RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR34475
ID AAR3
                                                                                                                                                                                                                                                                                                                                                                           AAR34474
* 8 × 0 0 0 0 0 0 × 8
                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                       AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vactines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other monounclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                  DNA and amino acid sequence of HCV envelope 1 and core proteins -used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                / Match 88.1%; Score 133; DB 17; Length 191; Local Similarity 92.3%; Pred. No. 2.9e-11; nes 24; Conservative 1; Mismatches 1: Indele
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV; non-A, non-B hepatitis virus; NANBHV; liver disease; polymerase chain reaction; diagnostic method.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNA and cDNA of hepatitis C virus - useful as probes diagnosing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Encoded by Hepatitis C virus clone JK3-A
                                                                                                                                                                      US DEPT HEALTH & HUMAN SERVICES.
US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPQRKTKRNTIRRPQDIKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (SANW ) SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                                                                        Claim 4; Page 209; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR34473 standard; Protein; 470
                                                                                                                                                                                                                               Purcell RH;
                                                                                                         95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91JP-0153736
                                                                                                                                        94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                               Bukh J, Miller RH,
                                                                                                                                                                                                                                                                WPI; 1996-139709/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1993-130638/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      191 AA;
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                                   N-PSDB; AAT16644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAQ40431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JP05068562-A.
                                 WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30-MAY-1991;
                                                                                                       15-AUG-1995;
                                                                                                                                        15-AUG-1994;
                                                                      22-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23-MAR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR34473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Seguence
                                                                                                                                                                         ( HSSU)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR34473
ID: AAR3
XX
AC AAR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
```

g

ò

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

ö

Gaps

ö

Indels

Length 470;

9

Page

```
09:20:19 2003
 æ
Fri Aug
```

```
The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against effection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, and antigen specificity of enzyme spectrum of immunoreactivity, and antigen specificity of enzyme immunosasays. This provides improved detection of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polymerase chain reaction; PCR; amplify; primer; non-A, non-B hepatitis; NANBH; virus; blood transmissible; detection; hepatitis virus; RT-PCR; C100 antibody; HCV RNA; NS5 region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                           New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                              Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                       85.4%; Score 129; DB 20; 92.3%; Pred. No. 1.4e-11; ive 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'label = Ser, Arg, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Len
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR53417 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Blood transmiscible NANBHV protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Key Location/Qualifiers
Misc-difference 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /label- Gln, Pro,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /label= Cys, Arg
Misc-difference 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             'label- Gly, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /label= His, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /label= Asp, Asn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              'label - Thr, Ala
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /label= Lys, Arg
Misc-difference 418
                                                                                                            Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Non-A, non-B hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 'label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
Best Local Similarity 92.3
Matches 24; Conservative
             WPI; 1999-204671/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 311
                                                                                                                                                                                                                                                                                                                                                                                       28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               17-JAN-1995
                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR53417;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR5341
ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-C. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 130; DB 14; Length 470;
Pred. No. 2e-10;
1; Mismatches 1; Indels
                                                                                             HCV; non-A, non-B hepatitis virus; NANBHV; liver disease; polymerase chain reaction; diagnostic method.
                                                                                                                                                                                                                                                                                                                                                                                       DNA and cDNA of hepatitis C virus – useful as probes for diagnosing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Indels
                                                              Encoded by Hepatitis C virus clone JK3-C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NC mosaic protein amino acid fragment C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (USSH ) US DEPT HEALTH & HUMAN SERVICES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPORKTKRNTIRRPODVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 4; Page 30-32; 44pp; Japanese.
                                                                                                                                                                                                                                                                                                    (SANW ) SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY06675 standard; Protein; 28 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       86.1%;
92.3%;
                                                                                                                                                                                                                                           91JP-0153736.
                                                                                                                                                                                                                                                                         91JP-0153736
                            30-JUL-1993 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 92.3
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                    WPI; 1993-130638/16,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sequences of these
See AAQ40425-Q40439
                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAQ40433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mosalc protein;
                                                                                                                                                                                                                                         30-MAY-1991;
                                                                                                                                                                                                                                                                      30-MAY-1991;
                                                                                                                                                                          JP05068562-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               17-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO9910506-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             04-MAR-1999
AAR34475;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Fields HA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY06675;
```

RESULT 12 AAY06675 ID AAY

ò 셤

```
expression cassette including a first polynuclectide region including a 5 non-coding region (NFN) sequence of an RNA virus and at least an N-terminal portion of a coding sequence of RNA virus, a second polynuclectide region including a 3 untranslated region (UTR) sequence of RNA virus, an second as the RNA virus and at least a C-terminal portion of a coding sequence of the virus and a third polynuclectide region encoding a reporter molecule, flanked by first and second polynuclectide regions; and a molecule, flanked by first and second polynuclectide regions; and a moner so as to enable a transcription of a minus strand RNA molecule from the expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette. Nucleic acid construct of the invention is useful for detecting the presence of an RNA virus in a cell. It is resistance of an RNA virus. The present sequence is Hepatitis C virus (HCV) isolate HCV-SI full-length polyprotein.
non-B hepatitis (NANBH) virus. The cDNA sequence was isolated using the primers given in AAQ63500-35. The amplified fragments are used in the detection of hepatitis virus. The target DNA was isolated from serum of chronically infected NANBH patients who were C100 antibody-positive and HCV RNA (NS5 region) positive. Reverse transcription-PCR and PCR were performed on CDNA and the total human NANBH DNA was constructed from 23 clones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              and a promoter operably linked RNA transcription of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              detecting the presence of RNA
                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nucleic acid construct; expression cassette; non-coding region; NCR; untranslated region; UTR; anti-viral drug; drug resistance; HCV-S1; Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The invention relates to nucleic acid construct which comprises an
                                                                                                                                                                                                                    ö
                                                                                                                                                                               Length 3010;
                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                  ä
                                                                                                                                                                                 Score 129; DB 15;
Pred. No. 2e-09;
                                                                                                                                                                                                                  0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Novel nucleic acid construct useful for virus, comprises an expression cassette to expression cassette for minus strand
                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (MOLE-) INST MOLECULAR & CELL BIOLOGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Example 1; Page 70-81; 81pp; English
                                                                                                                                                                                                                                                                                                                                                                     AAE20477 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hong
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV-S1 full-length polyprotein
                                                                                                                                                                                 85.4%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20-JUL-2001; 2001WO-IL00669.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24-JUL-2000; 2000US-220248P.
                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sg,
                                                                                                                                                                               Query Match 85.4
Best Local Similarity 92.3
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Lim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2002-280605/32.
N-PSDB; AAD33038.
                                                                                                                                                  3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EHRL/) EHRLICH G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ran YH, Lim SP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200208447-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                     01-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        31-JAN-2002
                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                      AAE20477;
                                                                                                                                                                                                                                                                                                                                   RESULT 14
                                                                                                                                                                                                                                                                                                                                                     AAE20477
                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                       8$$$$$$$$$$$
                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This sequence is encoded by the genome of a blood transmissible non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 - nsed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Blood-transmissible non-A non-B hepatitis virus DNA detection of hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 1; Page 8-20; 22pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                    Asp
                                                                                                                 Lys
                                                                                                                                                Val
                                                                                                                                                                                                                                                   Asp
                                                                                                                                                                                                                                                                                                                                                     Tyr
                                                                                                                                                                                                                                                                                                                                                                                                                       Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Thr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     label- Gly, Ile
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Met
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Lea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           G1y
             /label- Gly, Asp
                                                                                                                                                                                                                                                                                                                                                                                  label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     abel- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /label- Tyr,
                                                                                                                                                abel- Leu,
                                                                                                                                                                               label- Ser,
                                                                                                                                                                                                                  label- Thr,
                                                                                                                                                                                                                                                  label- Glu,
                                                                                                                                                                                                                                                                                                                                                   label- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                     .abel- Ala,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      abel- Glu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      abel- Val,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label- Lys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'label- Leu,
                                               label- Asn,
                                                                                                                                                                                                                                                                                 label- Met,
                                                                                                                                                                                                                                                                                                                  abel- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                    abel- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      label- Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          label- Ser,
                                                                                                               abel- Arg,
                                                                                label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92JP-0051885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92JP-0051885
                                                                                                                                                                                                                                                                                                 Misc-difference 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 1828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 1896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 2143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 2488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 2968
                                                                                                                                                                                               Misc-difference 1036
                                                                                                                                                                                                                                                                                                                                Misc-difference 1255
                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 1263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 1895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference 2989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 2862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1994-163130/20.
N-PSDB; AAQ63499.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (KAEN/) KAENNO K.
                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                             Misc-difference
                                                              Misc-difference
                                                                                              Misc-difference
                                                                                                                             Misc-difference
                                                                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JP06105690-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-APR-1994.
```

```
ö
                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The present sequence is that of an antigen used in a novel assay for determining anti-RCV antibody titre. A sample is diluted to an absorbance of not more than 2.0 and the peptide antigen is added. The method is useful for monitoring the effect of interferon treatment on anti-HCV antibody titres.
                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                           Ouery Match 85.4%; Score 129; DB 23; Length 3010; Best Local Similarity 92.3%; Pred. No. 2e-09; Matches 24; Conservative 0; Mismatches 2; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Determination of the antibody titre against hepatitis C virus antigen - using dilute solution and HCV antigenic peptide(s)
                                                                                                                                                                                                                                                                                                                                  HCV; non-A, non-B hepatitis virus; antigen; immunoassay;
interferon treatment; monitoring; antibody titre; detection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match

84.8%; Score 128; DB 16; Length 30;
Best Local Similarity 92.3%; Pred. No. 2.1e-11;
Matches 24; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                     Hepatitis C virus core antigenic peptide.
                                                                                                      1 PKPORKTKRNTIRRPODVKFPGGGVI 26
                                                                                                                    5 PKPQRKTKRNTYRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPORKTKRNTIRRPODVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Search completed: August 7, 2003, 11:14:09 Job time : 39.6364 secs
                                                                                                                                                                                                          AAR84558 standard; peptide; 30 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 1; Page 10; 11pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94JP-0082160.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94JP-0082160.
                                                                                                                                                                                                                                                                       06-JUN-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (ARIM/) ARIMA T.
(TORA ) TORAY IND INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1995-386063/50.
              3010 AA;
                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     16-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                            JP07260792-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                          13-OCT-1995.
              Sequence
                                                                                                                                                                                                                                          AAR84558;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                              RESULT 15
AAR84558
ID AAR84
XX
AC AAR84
XX
DT 06-JU
× og
                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                       SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

genome polyprotein genome polyprotein genome polyprotein polyprotein hepa polyprotein hepa genome polyprotein genome polyprotein genome polyprotein genome polyprotein hypothetical protein polyprotein polypeptide - hepa genome polyprotein genome polyprotein

\$21471 \$212707 \$010284 \$010925 \$010926 \$140711 \$18031 \$000881 \$000881 \$000881 \$000881 \$18030 \$18030

```
August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                          283308
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                     tal number of hits satisfying chosen parameters:
                                                                                                                                                                           US-09-491-146A-30
151
1 PKPORKTKRNTIRRPQDVKFPGGGVIYV 28
                                                                                                                                                                                                                                                                                                   283308 seqs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                     OM protein - protein search, using sw model
                                                                                                                                                                                                                                                BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                        Ainimum DB seq length: 0 Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR_76:*
1: pirl:*
2: pir2:*
3: pir3:*
4: pir4:*
                                                                                                                                                                           Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                       Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database
                                                                                                    Run on:
```

ave a	score greater than or equal to the score of the result being printed,	1
ance to h	sult being	ribution.
cted by ch	of the re	score dist
lts predic	the score	he total
r of resu	equal to	ysis of t
the numbe	than or	ed by anal
Pred. No. is the number of results predicted by chance to have a	ore greater	and is derived by analysis of the total score distribution.
Pre	သင္တ	anc

Description	qenome polyprotein				genome polyprotein																		genome polyprotein	genome polyprotein					
SUMMARIES	S41359	S41358	S41360 ·	PC2061	S41288	S41361	S19875	S41356	PC2060	A45573	S21336	S41353	S41355	S41357	S41348	S41371	S41341	S41370	S41369	S41368	S41342	S41344	S41350	S41354	S41345 .	S41347	4134	S41346	PQ0393
DB	7	~	7	7	~	7	7	~	~	н	7	~	7	~	~	7	7	7	~	7	7	7	~	~	~	7	~	~	~
a Query Match Length	114	114	124	411	Φ	123	782	108	411	3010	88	108	108	108	108	112	112	114	114	114	115	115	115	115	115	115	115	118	266
& Query Match	88.7	88.7	88.7	88.7	8	86.1	86.1	85.4	85.4	85.4	84.8	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1
Score	134	134	134	134	134	130	130	129	129	129	128	127	127	127	127	127	127	127	127	127	127	127	127	127	127	127	127	127	127
Result No.	п	7	m	4	Ŋ	9	7	œ	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29

RESULT 1  \$41359 genome polyprotein - hepatitis C virus (genotype 3, N2) (fragment) N; Contains: core protein C; Species: hepatitis C virus A; Variety; genotype 3, N2 C; Date: 19-Hay1994 sequence_revision 26-Jul-1996 *text_change 17-Nov-2000 C; Accession: \$41359 A; Accession: \$4135
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

ò g

```
protein; nonstructural protein; ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Genome polyprotein - hepatitis C virus (isolate JK3) (fragment)
N.Contains: core protein; envelope protein 1; nonstructural protein 2; NSI/E2 prote:
C,Species: hepatitis C virus
A,Variety: isolate JK3
C,Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C,Accession: S19875
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
Submitted to the EMBL Data Library, September 1991
A,Description: Sequence analysis of putative structural regions of Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Experimental source: isolate JK3
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
  C;Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N;Contains: core protein
C;Species: hepatLis C virus
A;Variety: genotype 3, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Molecule type: genomic RNA
A;Residues: 1-123 <VAN>
A;Cross-references: EMBL:229464; NID:9443890; PIDN:CAA82602.1; PID:9443891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Cross-references: EMBL: X61592; NID: 959482; PIDN: CAA43789.1; PID: 959483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                             Submitted to the EMBL Data Library, December 1993
A;Reference number: $41288
A;Accession: $41288
A;Accession: $41288
A;Accession: $41288
A;Accession: $41288
A;Coss_references: EMBL:X76918
C;Superfamily: hepatitis C virus genome polyprotein
C;Reyords: capid protein: ore protein; envelope protein; nonstr
C;Kryords: capid protein #status predicted <COR>
F;1-191/Product: core protein #status predicted <ENV>
F;192-372/Product: NS1 protein (fragment) #status predicted <NS1>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (genotype 3, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            рy
                                                                                                                                                                                                                                                                                                                                                                                                                    Length 492;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C) Accession: S41361
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5.
A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A. Experimental source: genotype 3, N4 C. Superfamily: hepatitis C virus genome polyprotein C. Keywords: capsid protein; core protein; polyprotein F.1-123/Product: core protein #status predicted <MAI>
                                                                                                                                                                                                                                                                                                                                                                                                                 Score 134; DB 2;
Pred. No. 1.7e-11;
0; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 130; DB 2;
Pred. No. 1.6e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPORKTKRNIVRRPODVRFPGGGOI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGVI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               86.1%;
88.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match
Best Local Similarity 96.2%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 86.1
Best Local Similarity 88.5
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Accession: $19875
A; Molecule type: genomic RNA
A; Residues: 1-782 <HON>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Reference number: S18029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: S41361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S19875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             841361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: mRNA
A; Residues: 1-411 <LLD3
A; Cross-references: 163:L12355; NID:g410169; PIDN:AAA20155.1; PID:g410170
C; Cross-references: GB:L12355; NID:g410169; PIDN:AAA20155.1; PID:g410170
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Reywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;
F; 192-383, Product: envelope protein E1 **status predicted <NPE>
F; 384-411, Product: nonstructural protein E2/NS1 **status predicted
F; 196, 209, 234, 305, 325/Binding site: carbohydrate (Asn) (covalent) **status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C;Accession: Pc2061
R;Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
Biochem. Biophys. Res. Commun. 199, 1474-1481, 1994
A:Title: Identification of the third major genotype of hepatitis C virus in France.
Reference number: Pc2060; MUID:94197744; PMID:8147893
                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                     A;Variety: genotype 3, N3
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C; Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Cross-references: EMBL:Z29463; NID:g443888; PIDN:CAA82601.1; PID:g443889
                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LIPA.
                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ó
                                                                                                                                                                                                                                                                        N3) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ρλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 124;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 411;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              'n
                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N;Contains: envelope protein El; nonstructural protein E2/NS1
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                         C; Accession: S41360
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Le: Call.
; Score 134; DB 2; Leng
....d No. 1.4e-11; I
1; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein - hepatitis C virus (fragment)
N;Contains: core protein; envelope protein; NS1 protein
C;Species: hepatitis C virus
Pred. No. 3.9e-12;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A:Experimental source: genotype 3, N3 C;Superfamily: hepatitis C virus genome polyprotein C;Reywords: capsid protein; core protein; polyprotein F;1-124/Product: core protein *status predicted <AMT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88.7%; Score 134; DB 2;
llarity 96.2%; Pred. No. 4.3e-12;
Conservative 0; Mismatches 1
                                                                                                                                                                                                                                                                        C virus (genotype 3,
                                                                                                                                30
                                                                               1 PKPORKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      #!!!!!!!!!!!!!!!!!!!!!!!!!!!!
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                            PKPQRKTKRNTIRRPQDVKFPGGGQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein N2 - hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88.7%;
96.2%;
Similarity 96.2%; 25; Conservative
                                                                                                                                                                                                                                                                           polyprotein - hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                            N;Contains: core protein
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           . Jecule type: genomic RNA A; Residues: 1-124 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Perence number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
ses 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
les 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cession: S41360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sccession: PC2061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loc
Matches
                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 5
S41288
                                                                                                                                                                                                                                                                           genome
                                                                                                                                                                                                                                            S41360
```

qq

à

g

ò

က

(nons

```
procein Naya; nonstructural protein Nab; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Species: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C;Accession: A45573
Naya A45573
Naya A5573
Naya A5573
Naya A5573
Naya A5573, NUD: 92295714; PMID: 1318627
Nayacession: A45573
Naya A5573, MUD: 92295714; PMID: 1318627
Nayacession: A45573
Naya A5573
Naya A5773
Naya 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Growe polyprotein S29 (core protein region) - hepatitis C virus (fragment) (5.5pecies: hepatitis C virus (5.5pecies: hepatitis C virus (5.5pecies: hepatitis C virus (5.5pecies: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 17-Nov-2000 C.5Accession: S21336
Bubmitted to the EMBL Data Library, April 1992
A.5Description: A sentitive serodiagnosis of hepatitis C virus infection with two A.5Reference number: S21336
A.5Accession: S21336
A.5Acc
                                                                                                                                 genome polyprotein - hepatitis C virus (strain JT)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98)
_protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   / Match 85.4%; Score 129; DB 1; Length 30 Local Similarity 92.3%; Pred. No. 5.6e-10; Nes 24; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      84.8%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
S41353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Accession: PC2060
R;Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
Bicohem. Blophys. Res. Commun. 199, 144-1481, 1994
A.Title: Identification of the third major genotype of hepatitis C virus in France.
A;Reference number: PC2060; MUID:94197744; PMID:8147893
A;Accession: PC2060
A;Access
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
F;1-191/Product: core protein #status predicted <MAT1>
F;192-383/Product: envelope protein 1 #status predicted <MAT2>
F;384-733/Product: NS1/E2 protein #status predicted <MAT3>
F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 2, N5
A; Variety: genotype 2, N5
Date: 19-May-1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ate: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Accession: 541356
Accession: S41356
Accession: Acces
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polyprotein - hepatitis C virus (genotype 2, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     85.4%; Score 129; DB 2; Length 411; 88.5%; Pred. No. 7.6e-11; ive 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                              2; Length 782;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 2; Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                             1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match 85.4%; Score 129; DB 2; Best Local Similarity 92.3%; Pred. No. 2e-11; Matches 24; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                 Score 130; DB 2;
Pred. No. 1e-10;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PKPQRKTKRNTCRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                        86.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                       24; Conservative
                                                                                                                                                                                                                                                                                                              Ouery Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
Matches 23; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

ö

ö

ö

ö

Gaps

ö

polyprotein - hepatitis C virus (genotype 2, N2) (fragment)

genome polyprotein - hepagenome N; Contains: core protein

```
Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
C;Species: hepatitis C virus
A;Variety: genotype 2, N2
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      %.Contains: core protein
C;Species: hepatitis C virus
A:Variety: genotype 2, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 2, N6
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          by LiPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
                                                                                                                                                                                                                                                            5 by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Rivan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994
A:Description: Analysis of hepatitis C virus genotypes 1 to 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                        Rivan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to A; Reference number: S41341
A; Reference number: S41353
A; Molecule type: genomic RNA
A; Residues: 1-108
A; Residues: 1-108
A; Cross-references: EMBL: Z29456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    84.1%; Score 127; DB 2; Le
92.3%; Pred. No. 3.8e-11;
No. Micmatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 127; DB 2; L
Pred. No. 3.8e-11;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A. Experimental source: genotype 2, N4
C. Superfamily: hepatitis C virus genome polyprotein
C. Keywords: capaid protein; core protein; polyprotein
F. 1-108/Product: core protein *status predicted <a href="Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A Experimental source: genotype 2, N2
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capaid protein; core protein; polyprotein
F;1-108/Product: core protein *status predicted AMT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPQRKTKRNINRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Cross-references: EMBL: 229460
A; Experimental source: genotype 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ouery Match
Best Local Similarity 92.3
Atches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match

Pest Local Similarity 92.37

- hea 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Reference number: S41341
                                                                                                                              C; Accession: S41353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Accession: 541355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C; Accession: $41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Accession: S41355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: S41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                   C;Species: hepatifis C virus
A;Variety: genotype 1, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            by LiPA
                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 108;
                                                                                                    Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C, Accession: $41348
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5
                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                        Score 127; DB 2; Le
Pred. No. 3.8e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match

84.1%; Score 127; DB 2; I
Best Local Similarity 92.3%; Pred. No. 3.8e-11;
Matches 24; Conservative 0; Mismatches 2;
C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: capsid protein; core protein; polyprotein F;l-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Experimental source: genotype 1, N6
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capaid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                   0; Mismatches
                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7, 2003, 11:21:48
                                                                                             Query Match
Best Local Similarity 92.3%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
A; Cross-references: EMBL: 229451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Search completed: August
Job time: 9.54545 secs
                                                                                                                                                                                                                                                                                                                                                                                                   N;Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: S41348
                                                                                                                                                                                                                                                                                                                      RESULT 15
S41348
```

us-09-491-146a-30.rsp

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

(without alignments) 268.226 Million cell updates/sec August 7, 2003, 11:05:41; Search time 4.90909 Seconds Run on:

US-09-491-146A-30

151 1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV 28 Perfect score: Sequence:

Scoring table:

Searched:

127863 seqs, 47026705 residues BLOSUM62 Gapop 10.0 , Gapext 0.5

tal number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

						•																												
	Description	9 h genome po	'n					ᆮ	၎	h genome	q	h genome	h genome		2	8 homo sapien							saccharon				_	8 caenorhabdi	3 xenopus lae		_	_	'n	1 clostridium
	Descr	000269	P27959	00140	00140	P2796	P2796	P2666	P2666	P2795	P2666	P2666	P2984	P2666	08036	P09958	P1971	04993	. P1812	Q4683	Q9ha6	Q8k9d	P0574	P2318	Q9er3	03298	02819	01940	P2157	00604	P95050	020063	P2152	P2413
	GI	POLG_HCVJT	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVTW	POLG_HCV1	RK4_TOBAC	FURI_HUMAN	POLG_BVDVN	RK4_SPIOL	CAT3_MAI2E	ACFD_ECOLI	FN3X_HUMAN	Y401_BUCAP	RL29_YEAST	FURI_MOUSE	FN3K_MOUSE	RL4_MYCLE	FURI_BOVIN	YSMK_CAEEL	YB1_XENLA	RL4_MYCBO	RL4_MYCTU	RK4_ARATH	FRA_DROME	BGAL_CLOAB
	DB	-	Н	-	-	Н	٦	-	_	Н	-	Н	~	Н	Н	-	-1	~	-	П	٦	-	Н	Н	-	-	Н	Н	٦	ч	-	-	-	-1
	Query Match Length	3010	513	3	N	m	737	3010	3010	3011	3033	3033	3010	3011	282	794	3988	293	496	1520	309	219	58	793	309	230	797	298	303	223	223	282	595	897
æ	Query	85.4	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1		78.8	35.1	34.4			32.5			31.8	31.1			30.5			30.1					29.8
	Score	129	127	127	127	127	127	127	127	127	127	127	120	119	53	52	50.5	49	49	4	48.5	48	47	4	46.5	46	4	45.5	45.5	45	45	45	45	45
	Result No.	1	7	m	4	Ŋ	9	۲.	æ	o	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	35	33

MEROPS; 229.001; ... MEROPS; U39.001; ... InterPro; IPR001410; DEAD.

PDB; 1A1Q; 25-MAR-98.

Q91713 xenopus lae	P47390 mycoplasma	P10637 mus musculu	P19332 rattus norv	Q10765 mycobacter1	P42921 bacillus su	Q9czx5 mus musculu	P28276 herpes simp	Q60769 mus musculu	P30594 rhizopus ol	Q8c2b3 mus musculu	P41128 brassica na
CHRD_XENLA	Y144_MYCGE	TAU_MOUSE	TAU_RAT	SYI_MYCTU	RL4_BACSU	PII1_MOUSE	IE63_HSV2H	TNP3_MOUSE	CHS1_RHIOL	HDA7_MOUSE	R131_BRANA
-	٦	-	Н	-	-	-	-	-	-	Н	н
941	279	732	751	1041	207	332	512	775	828	938	206
8.62	29.5	29.5	29.5	29.5	29.1	29.1	29.1	29.1	29.1	29.1	28.8
		Ŋ	'n	'n	44	44	44	44	44	44	٦. دى
4.5	44.5	44	44	44							4

## ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                     01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 26, Last sequence update)
01-APR-1993 (Rel. 42, Last annotation update)
03-APR-1993 (Rel. 42, Last annotation update)
04-APR-1993 (Rel. 42, Last annotation update)
05-APR-1994 (Mal); Protein El (GP32) (GP35); Envelope glycoprotein E2
05-APR-1996 (GP70) (Mal); Protein P7; Nonstructural protein NSA (P46); Nonstructural protein NSA (P47); Nonstructural protein NSA (P47); Nonstructural protein NSA (P46); Nonstructural protein NSA (P66); P70) (RNA-directed RNA polymerase) (EC 2.7.7.48);
05-APR-1993 (APR-1994); Nonstructural protein NSA (P66); Nonstructural protein NSA (P66); P70) (RNA-directed RNA polymerase) (EC 2.7.7.48);
05-APR-1993 (APR-1994); Nonstructural protein NSA (P66); P70) (RNA-directed RNA polymerase) (EC 2.7.7.48);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-92295714; PubMed-1318627;
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Nakazawa T., Hijkata M., Ishlmura Y., Shimotchno K.;
"Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among infected individuals.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VILTURE RES. 23:39-53(1992).

-1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBITC, SUGGESTING A POSSIBLE MEMBRANE-RELEATED FUNCTION.

NS3 AND NS5 MAY FLAY A ROLE IN THE VIRAL RNA REFLICATION.

-1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precision Polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (RNA)(N).
--- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
--- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                               3010 AA
                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D11168; BAA01943.1; -.
                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                              NCB1_TaxID=31642;
                                          POLG_HCVJT
Q00269;
                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
RESULT 1
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NSI)] (Fragment).
Hepatitis C virus (isolate HC-27) (HCV).
Viruses: SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPIIDASE.
CAPSID PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. MEDLINE-92230232; PubMed-1314459; MEDLINE-92230232; PubMed-1314459; Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro: IPR00252; HCV_capsid.
InterPro: IPR00252; HCV_capsid.
InterPro: IPR00251; HCV_core.
InterPro: IPR00251; HCV_core.
InterPro: IPR00251; HCV_NSI.
Pfam; PF0154; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01594; HCV_core; 1.
Pfam; PF01509; HCV_NSI; 1.
Probom; PF016602; HCV_NSI; 1.
Polyprotedin; Glycoprotein; Coat protein; Envelope protein; Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTÉNTIAL.

N-LINKED (GLCNAC.

N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
(GLCNAC.
      EMBL; D10074; BAA00968.1; -.
                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11111;
                                                                                                                               POLG_HCVJ2
P27959;
                                                                                                                                                                                                                                                                                                                                            Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INIT_MET
                                                                                 RESULT 2
POLG HCV12
ID I 01-AUG
ID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CELLUIAR AMINOPERIDAGE.
CAPSID PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS4/E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4/E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4/E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4/E (POTENTIAL).
POTENTIAL).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
ATP (POTENTIAL).
ATP (POTENTIAL).
N-LIKKED (GLANG...), (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 129; DB 1; Length 3010;
Pred. No. 1.1e-10;
0; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N - LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                     HCV_RdRP.
RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
;
      capsid
                                                               HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4a.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                    PF01538; HCV_NS2; 1.

PF02907; HCV_NS3; 1.

PF01006; HCV_NS4s; 1.

PF01001; HCV_NS4s; 1.

PF01506; HCV_NS3s; 1.

PF00271; helicase_C; 1.

PF00998; Viral_RGRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Match 85.4%;
Local Similarity 92.3%;
tes 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  326573
                                                                                                                                                                                                                                                Pfam; PP01543; HCV_caps1d; Pfam; PF01542; HCV_core; 1. Pfam; PF01539; HCV_env; 1. Pfam; PF01560; HCV_NSI; 1.
IPRO02522; H
IPRO02521; H
IPRO02519; H
IPRO02531; H
IPRO02518; H
                                                                                                         IPR004109;
IPR000745;
                                                                                                                                                                                                             IPR007095;
                                                                                                                                                                                                                                 IPR007094;
                                                                                                                                                  IPR001490;
                                                                                                                                                                    IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2041
2077
2240
2529
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF02907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF00998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3D-structure.
INIT_MET
                                         InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
Matches 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ProDom;
SMART; S
                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
```

PKPORKTKRNTIRRPQDVKFPGGGVI 26

(POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL).

N-LINKED

(POTENTIAL). (POTENTIAL). (POTENTIAL). (POTENTIAL).

So

g

```
84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; D10687; BAA01529.1; -.
                                                                  Query Match 84.19
Best Local Similarity 92.33
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1115

> 383

369

1369

209

233

418
      520 5
520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-31644;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Gl
Transmembrane;
                                                                                                                                                                                                                                                                                          POLG_HCVHK
Q01403;
    NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                       ACSULT 4
POLG_HCVHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
      FT
SO
                                                                                                                                                  ò
                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                THE THEFFE TEFFE T
                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Biolinformatics and the EMBL outstation-the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license alreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL)
BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                             01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope 91ycoprotein El (GP32) (GP35); Envelope 91ycoprotein El (GP68) (GP60) (NS1)) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatítis C virus (isolate HCV-476) (HCV).
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                    ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                             Length 513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN ( POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein; Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
                       Score 127; DB 1; Length 51
Pred. No. 3.3e-11;
Pred. No. 3.3e-11; Indels
  943F31E3514CDEF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                        520 AA
                                                                                 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (
N-LINKED (
N-LINKED (
N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                          1 PKPORKTKRNTIRRPODVKFPGGGVI 26
                                                                                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; D10688; BAA01530.1; ...
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01549; HCV_core; 1.
Pfam; PF01559; HCV_mon; 1.
Pfam; PF01559; HCV_mon; 1.
55704 MW;
                                        ch 84.1%;
1 Similarity 92.3%;
24; Conservative
                                                                                                                                                                                                                                                                        STANDARD;
513 AA;
                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus
                                                                                                                                                                                                                                                                      POLG_HCVH4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INIT_MET
                                                                                                                                                                                                                           RESULT 3
POLG_HCVH4
ID POLG_H
AC 001404
DT 01-JUL
DT 16-OCT
DE Genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
                                                                                    Matches
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isp-sib.ch)
                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLULAR AMINOPEFIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (FOTENTIAL).
MAJOR ENVELOPE PFOTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genome polyprotein (Contains: Capsid protein (Core protein) (P22); Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment). Hepatitis C virus (isolate HCV-KF) (HCV). Wiruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-93019030; PubMed-1383400;
ADE K., Inchauspe G., Fujisawa K.;
ADE K., Inchauspe G., Fujisawa K.;
ADE K., Inchauspe G., Fujisawa K.;
AGENOMIC characterization and mutation rate of hepatitis C virus
1solated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN BAYBLOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nonstructural protein.
                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR: JQ1925; JQ1925.
InterPro: IPR002521; HCV_capsid.
InterPro: IPR002521; HCV_core.
InterPro: IPR002519; HCV_core.
InterPro: IPR002519; HCV_env.
InterPro: IPR002519; HCV_env.
InterPro: IPR002519; HCV_env.
InterPro: IPR002519; HCV_core: 1.
Pfam; PF01542; HCV_core: 1.
Pfam; PF01542; HCV_env; 1.
Pfam; PF01550; HCV_env; 1.
ProDom; PD186062; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                                              Length 520;
                                                               Score 127; DB 1; Length 52. Pred. No. 3.4e-11;
520
56499 MW; AA135246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BY SIMILARITY.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                     520 AA
                                                                                                                                                     0; Mismatches
                                                                                                                                                                                                                                       5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                         1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-20001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation); Matrix
16-OCT-2001 (Core protein); Matrix
17-OCT-2001 (Core protein); Matrix
17-OCT-2001 (Core protein); Matrix
17-OCT-2001 (Core protein); Matrix
18-OCT-2001 (Core protein); Matrix
18-OCT-2001 (Core protein); Matrix
18-OCT-2001 (Rel. 40, Matrix
18-OCT-2001 (Rel. 40,
                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -i- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION. NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-i- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPMCTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MANA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-92230232; PubMed-1314459; Okado M., Lizuka H., Okada S.I., Yamamoto K., Lizuka H., Cokada S.I., Okada S.I., Tanaka T., Fukuda S., Tsuda F., Mishiro S., Full-length sequence of a hepertitis C virus genome having poor monology to reported isolates: comparative study of four distinct Virology 188:331-341(1992).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL, D10075; BAA00969.1; -.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR02531; HCV_env.
IPR02531; HCV_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Transmembrane; Nonstructural protein.
INIT MET 1 REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                              Length 520
                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                     Score 127; DB 1; Le
Pred. No. 3.4e-11;
                                                                                                                                                              56476 MW; 1D2BD0A6FF27349B
                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                          84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
424
431
449
520
520 AA;
                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SECUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11112;
                                                                                                                                                                                                                                                                                                                         24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVJ5
P27960;
                                                                                                                         NON_TER
SEQUENCE
                                         CARBOHYD
                                                                               CARBOHYD
                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVJS
ID POLG_H
AC P27960
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib.ch).
                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last sequence update)
Genome polyprotein (Contains annotation update)
Genome polyprotein (Contains Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NSI and NS2] (Fragment).
Heparitis C virus (isolate HC-J7) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genotypes.;

Virology 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
- FUNCTION: THE SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCARSID COVERED BY LIPOPROTEIN MAND GLYCOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-92230232; PubMed-1314459; Okamanoto K., Lizuka H., Cokamoto H., Kurai K., Okada S.I., Yamanoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                     (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                           POTENTIAL).
                                                                                                                                                                                                                                                                              Score 127; DB 1; Length 737, Pred. No. 5e-11;
                                                                                                                                                                                                                                                    81207 MW; 3AF699D82AD501B1 CRC64;
             N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                              Ź
                                                                                                                                                                                                                                                                                                           0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                              737
                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_cor.
InterPro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                              th 84.1%;
Similarity 92.3%;
24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D10077; BAA00971.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                           POLG_HCVJ7
P27961;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
NON_TER
SEQUENCE
               CARBOHYD
CARBOHYD
                                           CARBOHYD
CARBOHYD
                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                                                                                                                                                                                                                    a
```

us-09-491-146a-30.rsp

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-502 (Rel. 2), Last annotation update)
15-5EP-2003 (Rel. 42, Last annotation update)
16-5EP-2003 (Rel. 42, Last annotation update)
16-5EP-2003 (Rel. 42, Last annotation update)
16-5EP-2003 (Rel. 42, Last annotation update)
16-5ED-2003 (Rel. 42, Last annotation Elevation Elevator 
                                                                                                                                                         CELLUIAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
NONSTRUCTURAL PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ., Okayama H.;
the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; "Structure and organization of the hepatitis C virus genome isolate
                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEGUENCE OF 1487-1500.
MEDILTB-962324; Pubmed-8647104;
BOTOWSKI P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  84.1%; Score 127; DB 1; Length 737; 92.3%; Pred. No. 5e-11;
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Mismatches
                                                                                                                                                                                                                                                                                                   POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPORKTKRNTIRRPODVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-91140698; Pubmed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity 92.3
les 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       from human carriers."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-AUG-1992 (Rel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                                       INIT_MET
                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                              TRANSMEM
                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 7
POLG_HCVBK
ID POLG_H
AC P26663
DT 01-AUG
                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the SWISS Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                              X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.

MEDLINE-99227646; PubMed-9568991;

X and Y., L1 Y., Munshi S., Sardana W., Cole J.L., Sardana M.,
Steinkuehler C., Tomel L., de Francesco R., Kuo L.C., Chen Z.;
"Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
" virus: a 2.2-A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-647(1998).

-1 FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELACATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL FNA REPLICATION.

-1 CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in P1 and Ser or Ala in P1'.

-1 CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [RNA](N).
-I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MINA.
                                                                         MEDLINE-97015088; PubMed-8861916;
LOVE R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
"The crystal structure of hepatitis C virus NS3 proteinase reveals
trypain-like fold and a structural zinc binding site.";
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RNA_pol_DS_PS
RNA_pol_PSvir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF01543; HCV_capsid;
PF01542; HCV_core; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M58335; AAA72945.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10UV; 26-JUN-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001490; H
InterPro; IPR002868; H
InterPro; IPR002166; H
InterPro; IPR007095; R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002531; FinterPro; IPR002518; FinterPro; IPR004109; FinterPro; IPR000745; F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A38465; GNWVTC.
1A1Q; 25-MAR-98.
1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               nterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEROPS; U39.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1087;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1NS3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PDB;
```

PF01560; PF01538; 1 PF01006;

Ptam;

```
POLG_HCVJA
P26662;
                                                                        SEQUENCE
    TURN
STRAND
                                             TURN
STRAND
                                                                                                                                                                                                                           RESULT 8
POLG_HCVJA
                                 HELIX
                                                                                                                                                                                                                                                                       555558
                                                                                                                                                                                 a
                                                                                                                                                           ò
                                                                                                                                        CELLULAR AMINOPEPTIDASE.
CAPESID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.1.EZ (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01006; HCV_NS5a; 1.
Pfam; PF01066; HCV_NS5a; 1.
Probom; PD186062; HCV_NS1; 1.
SMART: SM0487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Cat protein; Envelope protein; Holicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                              REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                 POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                        ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1161
1163
1166
1171
1171
                                                                                                                                                                                                  729
1006
1615
1615
1862
2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 081
                                                                                                                                                                                                                                                                                 347
1083
1107
                                                                                                                                                                                                                                                                                                                        1165
1165
1316
196
209
234
250
305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10331
10339
1059 0
1075 9
1075 1
1086 1
1087 1
1087 1
1087 1
1088 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098 1
1098
                                                                                                               3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                           ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                 ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                  TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TURN
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAND
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TURN
STRAND
TURN
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
STRAND
STRAND
STRAND
TURN
STRAND
STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                          CHAIN
                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HELIX
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last sequence update)
01-AUG-1992 (Rel. 41, Last annoctation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3.(P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Japanese) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROPEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISCUSSION OF SEQUENCE.
MEDLINE-91192160; PubMed-1849488;
Kato N. Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K.,
Ohkoshi S., Shimotohno K.;
Molecular structure of the Japanese hepatitis C viral genome.";
FEBS Lett. 280:325-328(1991).
                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
MEDIJERS-9108550; PubMed-2175903;
KATO N., H11/KATE M., OOGSUYAMA Y., NAKAGAWA M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
"Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                 Length 3010;
                                                                                                                                                                                                                              Score 127; DB 1; Length 30 Pred. No. 2.3e-10; 0; Mismatches 2; Indels
                                                                                                                                                                   327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                        PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                           Query Match
Best Local Similarity 92...
Best Cocal Similarity 92...
"hea 24; Conservative 0
                                                                                                                                                                                                                                     84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
1188
1197
1202
1204
1688
1187 118
1189 119
1198 120
1203 120
1680 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (RNA)(N).
```

PKPQRKTKRNTIRRPQDVKFPGGGVI

٦

```
. d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET 1 CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 127; DB 1; Length 3u. Pred. No. 2.3e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA993794F46DB185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LINKED
                                                                   ### PRO0252; HCV_capsid.
### Proc. IPR00252; HCV_capsid.
### PRO0253; HCV_care.
### PRO0253; HCV_env.
### PRO0253; HCV_env.
### PRO0253; HCV_NS1.
### PRO0353; HCV_NS3.
### PRO0350; HCV_NS4.
### PRO0350; HCV_NS4.
### PRO0350; HCV_NS4.
### PRO0350; HCV_NS4.
### PRO0350; HCV_NS5.
### PRO0350; HCV_CAPS.
##
                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 84.1
Best Local Similarity 92.3
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SMART; SM00487; DEXDC;
                                                 InterPro; IPR001410;
InterPro; IPR002522;
                                                                                        InterPro; IPR002521;
                                                                                                         InterPro; IPR002519;
                                                                                                                            InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA;
                MEROPS; S29.001; -.
                                                                                                                                                              InterPro; I
InterPro; I
                                                                                                                                                                                                    InterPro;
                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                           InterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                              Pfam;
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institution. There are no restrictions on its use by non-profit institutions and also so its content is no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                  01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annoctation update)
6come polyprotein (Contains: Capsid protein C (Core protein) (P22);
6come polyprotein E1 (GP32) (GP35); Envelope glycoprotein E2
6GP69) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
6C 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural pr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hoop Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          oligonucleotide: the crystal structure provides insights into the mode
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -I-FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3.
-I-FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF
NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NSB.
-I-FUNCTION: NS4A-NS4B, NS4B-NS5A AND NS5A-NSB.
-I-FUNCTION: NS5A SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-I-FUNCTION: NS5B IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN
ESSENTIAL ROLE IN THE VIROS REPLICATION.
-I-CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
position, Cys or Thr in Pl and Ser or Ala in Pl'.
-I-CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -i- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
-i- PTM: THE STRUCTURAL PROTEINS C. E1 AND E2 ARE PRODUCED BY PROTECLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
-i- SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY U39.
-i- SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates.";
Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitan: M., Nasoff M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657
MEDLINE-98154321; Pubmed-9493270;
PKPQRKTKRNTNRRPQDVKFPGGGQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          of unwinding.";
Structure 6:89-100(1998).
                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus
                                                                                                                                                                                              POLG_HCVH
P27958;
                           'n
                                                                                                                            RESULT 9
POLG_HCVH
```

ö

Gaps

us-09-491-146a-30.rsp

```
Gaps
          N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 127; DB 1; I
Pred. No. 2.3e-10;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1411
1417
1420
1420
1436
11453
11463
1480
1480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1627 162
1635 163
1640 165
3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                             24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
POLG_HCVJ6
ID POLG_HCVJ6
      CARBOHYD
CARBOHYD
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HELIX
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                   TURN
TURN
HELIX
TURN
STRAND
TURN
STRAND
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                    STRAND
TURN
HELIX
TURN
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
STRAND
STRAND
TURN
STRAND
HELIX
STRAND
STRAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Loc
Matches
          g
        (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ProDom; PD186062; HCV_NS1; 1.
SMART; SMO4BY; DEXDC; 1.
PO1yprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
3D-structure. 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
CHARGE RELAY SYSTEM (BY SIMIL CHARGE RELAY SYSTEM (BY SIMIL CHARGE RELAY SYSTEM (BY SIMIL ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC...) (POTEN N-LINKED (GLCNAC...)) (POTEN N-LINKED (GLCNAC...) (POTEN N-LINKED (GLCNAC...)) (POTEN N-LIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROTEIN P7
entities requires a license agreement (
or send an email to license@lsb-sib.ch)
                                                                                                                                                                                                                                                                      InterPro; IPR002519; HCV_env.
InterPro; IPR002511; HCV_env.
InterPro; IPR002511; HCV_NS1.
InterPro; IPR001210; HCV_NS2.
InterPro; IPR001409; HCV_NS2.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001266; HCV_NS4b.
InterPro; IPR001666; HCV_RGRP.
InterPro; IPR001656; HCV_RGRP.
InterPro; IPR001095; RNA_POL_DS_PS.
InterPro; IPR001099; RNA_POL_DS_PS.
InterPro; IPR001099; RNA_POL_PS_PS.
InterPro; IPR001099; HCV_env; I.
Pfam; PF01543; HCV_env; I.
Pfam; PF01543; HCV_env; I.
Pfam; PF01543; HCV_env; I.
Pfam; PF01500; HCV_NS1; I.
Pfam; PF01006; HCV_NS1; I.
Pfam; PF01001; HCV_NS4b; I.
Pfam; PF01001; HCV_NS4b; I.
Pfam; PF01001; HCV_NS4b; I.
Pfam; PF01001; HCV_NS4b; I.
Pfam; PF01006; HCV_NS5a; I.
Pfam; PF00998; Viral_RGRP; I.
Proxon; PD186062; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
                                                                                                                                                                                                                                         HCV_capsid.
HCV_core.
                                                      EMBL; M67463; AAA45534.1; -. PIR; A36814; GNWVCH. PDB; 1HEI; 25-NOV-98. PDB; 1A1V; 16-FEB-99. PDB; 1A1V; 17-JUN-98. MEROPS; 529.001; -. TRANSFAC; T04155; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPRO01410; D
IPRO02522; H
IPRO02521; H
IPRO02511; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1192

1384

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

11052

1105
                                                                                                                                                                                                                                       InterPro;
InterPro;
                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

```
PDIG_HCVJ8 STANDARD; PRT; 3033 AA.

POLG_HCVJ8 STANDARD; PRT; 3033 AA.

POLG_HCVJ8 STANDARD; PRT; 3033 AA.

O1-AUG-1992 (Rel. 23, Created)

T 28-FEB-2003 (Rel. 23, Last sequence update)

T 28-FEB-2003 (Rel. 41, Last annotation update)

E Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);

E Greboloe glycoprotein E1 (GP35); Envelope glycoprotein E2

E GF68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)

E EC 3.4.22.-); Protease-Ahelicase NS3 (P70) (Heperivitin)

E EC 3.4.22.-); Protease-Ahelicase NS3 (P70) (Heperivitin)

E NS4B (P27); Nonstructural protein NS4A (P4); Nonstructural protein

ENS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein

E NS5B (P66) (P70) (RNA-directed RNA Polymerase) (EC 2.7.7.48)].

E Hepatitis C virus (isolate HC-43) (HV4).

Viruses; SSRNA Positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                              Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
W Polyprotein; Glycoprotein; Envelope protein; Helicase; ATP-binding;
W Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
T INIT_MET 1 SEMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
T CHAIN 1 115 CAPSID PROTEIN C (POTEWITAL).
T CHAIN 192 383 MAJOR ENVELOPE PROTEIN NSI (POTENTIAL).
T CHAIN 734 1010 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
T CHAIN 1010 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
T CHAIN 1011 1619 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
T CHAIN 1620 1866 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
T CHAIN 1867 2017 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                  RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MW; F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 127; DB 1;
Pred. No. 2.3e-10;
0; Mismatches 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             649 649
1091 1091
2038 2038
2811 2811
3033 AA; 329165 M
Pfam; PF01506; HCV_NSSa; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NSI; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                             11111
1169
1241
1323
196
209
234
305
                                                                                     SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus.
NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 11
POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                 MEDILINE-9204440; PubMed-1658196;
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
Machida A., Miyakawa Y., Mayumi M.;
Mucheotide sequence of the genomic RNA of hepatitis C virus isolated
from a human carrier: comparison with reported isolates for conserved
T and divergent regions. "I sequence of the genomic RNA AND NSA AND NSA AND NSA PROFILE J. Gen. Virol. 72:8697-2704(1991).
I. Gen. Virol. 72:8697-2704(1991).
I. FUNCTION: THE SMALL PROFILE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
OF CHARITIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precusor polyprotein, commonly with Asp or Glu in the P6
posttion, Cys or Thr in Pl and Ser or Ala in Pl'.
C -1 CATALTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
          01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 41, Last annoctation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21) (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS58 (P56); Nonstructural protein NS58 (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
NS58 (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Nituses: SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (RNA)(N).

SUBURIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Interpro; IPR002521; HCV_core.
Interpro; IPR002519; HCV_env.
Interpro; IPR002511; HCV_NS1.
Interpro; IPR00518; HCV_NS2.
Interpro; IPR004109; HCV_NS3.
Interpro; IPR004409; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR001490; HCV_NS4b.
IPR002868; HCV_NS5a.
IPR002166; HCV_RGRP.
IPR001650; Helicase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D00944; BAA00792.1; -. PIR; JQ1303; JQ1303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HSSP; P27958; IHEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                          NCBI_TaxID=11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF02907;
PF01006;
PF01001;
                                                                                                                                                                                                                                                                       Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             <u>-</u>
```

ö

Gaps

```
MEDLINE-92230206; Pubmed-1314449;
Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
The Taiwanese hepatitis C virus genome: sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA.";
Virology 188:102-113(1992).
I- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HUDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
I- CATALITIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
          RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                       SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               position, Cys or Thr in Pl and Ser or Ala in Pl'. CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate
                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein, commonly with Asp or Glu in the P6 Cys or Thr in P1 and Ser or Ala in P1'.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 127; DB 1; Length 3033;
Pred. No. 2.3e-10;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                     .1 N-LINKED (GLCNAC. . .) (POT 330177 MW; 1A173E7E3381FD1A CRC64;
                                       (BY
(BY
                                       CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                 N-LINKED GLCNAC
                                                                                       ATP (POTENTIAL).
DECH BOX.
                                                                                                                       N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sh 84.1%;
| Similarity 92.3%;
24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                1091 109
2038 203
2359 235
2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVTW
P29846;
                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
      CHAIN
TRANSMEM
ACT_SITE
                                                                  ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 12
POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best
      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CELULAR ANIOPEPTIDES.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                               [RNA](N)
-!-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPMOTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MANA.
SEQUENCE FROM N.A. MEDLINE-92230232; PubMed-1314459; Okamanoto K., Lizuka H., Tanaka T. Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T. Eukuda S., Tsuda F., Mishiro S.; Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR007095; RNA_pol_DS_PS. IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR004109; HCV_NS3.
IPR000745; HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_RdRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_capsid; 1.
HCV_core; 1.
HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWJB. HSSP; P27981; HET. MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002531; HCV_NS1
InterPro; IPR002518; HCV_NS2
InterPro; IPR004109; HCV_NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 interPro; IPR002519; HCV_env
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS1; 1.
HCV_NS2; 1.
HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SM00487; DEXDC; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .nterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              interPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
```

, was a same a same

ö

Gaps

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Last annotation Elementary (GP6) (NS1); Proteansex/hellosse NS3 (PR0) (Hepacivitin)
16-SEP-2003 (NS1); Proteansex/hellosse NS3 (PR0) (Hepacivitin)
16-SEP-2003 (NS1); Nonstructural protein NS5A (PS6); NONSTRUCTURAL NS5A (PS6); NONSTRUCTUR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                 (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                            Length 3010;
430 N-LINKED (GLCNAC. . .) (POTENT 448 N-LINKED (GLCNAC. . .) (POTENT 540 N-LINKED (GLCNAC. . .) (POTENT 540 N-LINKED (GLCNAC. . .) (POTENT 556 N-LINKED (GLCNAC. . .) (POTENT 623 N-LINKED (GLCNAC. . .) (POTENT 623 N-LINKED (GLCNAC. . .) (POTENT 645 N-LINKED (GLCNAC. . .) (POTENT 2071 N-LINKED (GLCNAC. . .) (POTENT 2077 N-LINKED (GLCNAC. . .) (POTENT 2077 N-LINKED (GLCNAC. . .) (POTENT 2240 N-LINKED (GLCNAC. . .) (POTENT 2240 N-LINKED (GLCNAC. . .) (POTENT 2788 N-LINKED (GLCNAC. . .) (POTENT 27
                                                                                                                                                                                                                                                                                                                                    Score 120; DB 1; Lengtn JU. Pred, No. 2.5e-09;
                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6 KPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 KPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                       79.5%;
92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                        23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                AA;
                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=11104;
                                                                                                                                                                                                                                                                                                           3010
                                                                       540
556
576
623
623
2041
22240
2529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 [RNA](N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCV1
                                                                    CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                        CARBOHYD
       CARBOHYD
                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                     Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P26664;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 13
POLG_HCV1
                                                                                                                                                                                                                                                                                                                                                                                                        Matches
       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute of There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.

GORE PROTEIN (POTENTIAL).

MAJOR ENVELOPE PROTEIN E, (POTENTIAL).

MAJOR ENVELOPE PROTEIN NSI, (POTENTIAL).

NONSTRUCTURAL PROTEIN NSI, (POTENTIAL).

PROTEASE, HELICASE NSI (POTENTIAL).

NONSTRUCTURAL PROTEIN NSA (POTENTIAL).

NONSTRUCTURAL PROTEIN NSA (POTENTIAL).

NONSTRUCTURAL PROTEIN NSA (POTENTIAL).

RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                            SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ..us; HCV_NS5a.
..rrC; IPR007166; HCV_RGRP.
.trrC; IPR007095; RNA_pol_DS_PS.
.tr.Pro; IPR007094; RNA_pol_PSvir.
'm; PF01543; HCV_capsid; 1.
'; PF01542; HCV_core; 1.
PF01566; iff.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                EMBL; M84754; -; NOT_ANNOTATED_CDS.
PIR; A40244; GNWVTW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RGRE; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_env;
HCV_NS1;
HCV_NS2;
HCV_NS3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002518;
                                                                                                                                                                                                                                                                                                                               1N64; 25-FEB-03
1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2014
347
11083
1107
11165
1230
1316
1316
209
233
234
250
417
417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01006; PF01001; P
                                                                                                                                                                                                                                                                                                                                                                             MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PF02907
                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      nterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               etam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     stam;
```

ô

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                              SERAIN-CV. Bright Yellow 4; TISSUE-Leaf;
STRAIN-CV. Bright Yellow 4; TISSUE-Leaf;
STRAIN-CV. Bright Yellow 4; TISSUE-Leaf;
STOACCO Chloroplast ribosomal protein L4.";
Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
-- FUNCTION: THIS PROTEIN BINDS DIRECTLY AND SPECIFICALLY TO 23S RRNA
(EX.SIMILARITY). MAY PLAY A ROLE IN PLASTID TRANSCRIPTIONAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                               Gaps
                                                                                                                                                                                                                                                                                                                     Nicotiana tabacum (Common tobacco).

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Bermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REGULATION.
--- SUBCELLULAR LOCATION: Chloroplast.
--- SIMILARITY: BELONGS TO THE L4P FAMILY OF RIBOSOMAL PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FURI_HUMAN STANDARD; PRT; 794 AA. P09958; 014336; 01-MAR-1989 (Rel. 10, Created) 01-MAR-1990 (Rel. 14, Last sequence update) 15-SEP-2003 (Rel. 42, Last annotation update) Furin precursor (EC 3.4.21.75) (Paired basic amino acid residue cleaving enzyme) (PACE) (Dibasic processing enzyme).
                                                                                                                                                                                                                                     30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
50S ribosomal protein L4, chloroplast precursor (R-protein L4).
                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002136: Ribosomal_L4/L1E.
Pfam; PF00573; Ribosomal_L4: 1.
Ribosomal protein; rRNA-binding; Chloroplast; Transit peptide.
Ribosomal protein; rRNA-binding; Chloroplast; Transit peptide.
TRANSIT 1 43 CHLOROPLAST (BY SIMILARITY).
CHAIN 44 282 508 RIBOSOMAL PROTEIN L4.
DOMAIN 254 282 ASP/GLU-RICH (ACIDIC).
SEQUENCE 282 AA; 31252 MW; 12BFAC9C08D9E4AC CRC64;
                              Score 119; DB 1; Length 3011;
Pred. No. 3.5e-09;
1; Mismatches 3; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 53; DB 1; Length 282; Pred. No. 1.7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
 327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        . 9
                                                                                                                                                                                                            282 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. 1.7;
5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PKPQRKT --- KRNTIRRPQDVKFPGGGVIY 27
                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                EMBL; AB010878; BAA31510.1; -. PIR; T01739; T01739.
                               78.8%;
84.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ch 35.1%;
1 Similarity 43.3%;
13; Conservative
                                             Local Similarity 84.6 les 22; Conservative
                                                                                                                                                                                                            STANDARD;
 3011 AA;
                                                                                                                                                                                                          RK4_TOBAC
080361;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    108
SEQUENCE
                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 15
FURI_HUMAN
                                                                                                                                                                            RESULT 14
RK4_TOBAC
                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                           g
                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    884444B
                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELLULAR ANINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR WAJOR PROTEIN (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
CHARGE RELAY SYSTEM (I
CHARGE RELAY SYSTEM (I
CHARGE RELAY SYSTEM (I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                             LICEPPO; IPR002166; HCV_NS5a.
LICEPPO; IPR001656; HCV_RGRP.
INTERPO; IPR007095; RNA_POl_DS_PS.
INTERPO; IPR007094; RNA_POl_DS_PS.
INTERPO; IPR007094; RNA_POl_DS_PS.
INTERPO; IPR007094; RNA_POl_PSV1r.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01559; HCV_env; 1.
'fam; PF01550; HCV_NS1; 1.
'fam; PF01550; HCV_NS1; 1.
'fam; PF01550; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                             interPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
                                                                       InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                          PF01538; HCV_NS2; 1.
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS5s; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00999; Viral_RGRP; 1.
Probom; PD160662; HCV_NS1; 1.
                                                                                                                          InterPro; IPR002519; HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     369
11083
11107
11165
1237
1237
1237
1237
423
423
4430
540
540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              645
2041
2077
2240
2364
2789
           PDB; 1A1V; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
                                             MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
```

ä

the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).

EMBL; X17094; CAA34948.1; -. EMBL; X04329; CAA27860.1; -. EMBL; A06939; CAA00605.1; -. PIR; A39552; KXHUF. HSSP; Q99405; 1MPT.

Genew; HGNC:8568; FURIN.

MIM; 136950;

MEROPS; S08.071; -

13

Page

```
SEQUENCE OF 296-794 FROM N.A.
MEDLINE-87053858; PubMed-3023061;
Roebroek A.J.M., Schalken J.A., Leunissen J.A.M., Onnekink C.,
Bloemers H.P.J., van de Ven W.J.M.;
"Evolutionary conserved close linkage of the c-fes/fps proto-oncogene and genetic sequences encoding a receptor-like protein.";
EMBO J. 5:2197-2202(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                     14)
3D-STRUCTURE MODELING OF CATALYTIC DOMAIN.
MEDITINE-942916.19; PubMed-8020465;
Siezen R.J., Creemers J.W.M., van de Ven W.J.M.;
"Homology modelling of the catalytic domain of human furin. A model for the eukaryotic subtilisin-11ke proprotein convertases.";
Eur. J. Blochem. 222:255-266(1994).
                                                                                                                                                                                                                                                                                  "CDNA and gene structure for a human subtilisin-like protease with cleavage specificity for paired basic amino acid residues."; DNA Cell Biol. 10:319-328(1991).
                             Craniata; Vertebrata; Euteleostomi;
Catarrhini; Hominidae; Homo.
                                                                                                  TISSUE-Blood;
MEDLINE-90175002; PubMed-2408021;
van den Ouweland A.M.W., van Duljnhoven H.L.P., Keizer G.D.,
Dorssers L.C.J., van de Ven W.J.M.;
"Structural homology between the human fur gene product and the
subtilisin-like protease encoded by yeast KEX2.";
Nucleic Acids Res. 18:664-664(1990).
                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE-91321735; PubMed-1713771;
Barr P.J., Mason O.B., Landsberg K.E., Wong P.A., Klefer M.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-92332543; PubMed-1629222;
              Homo sapiens (Human).
Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
OR PACE OR FUR
                                                                                     SEQUENCE FROM N.A.
                                                         NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                        Brake A.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROCESSING
```

```
ö
     Score 52; DB 1; Length 794; Pred. No. 7.3; 4; Mismatches 5; Indels
                                                                                                                                                                                                                                            7, 2003, 11:20:03
                                                                                                                                          4 ORKTKRNTIRRPODVKFP 21
     34.4%;
50.0%;
                                                           Conservative
                                                                                                                                                                                                                                            completed: August
                                Local Similarity
                                                           6
        Query Match
                                                                                                                                                                                                                                            Search com
Job time :
                                                           Matches
                                                                                                               8
                                                                                                                                                          g
                                                                                               SUBCELLULAR LOCATION: SEEMS TO BE LOCALIZED INTRACELLULARLY TO THE TRANS GOLGI NETWORK. PROPEPTIDE CLEAVAGE IS A PREREQUISITE FOR EXIT OF FURIN MOLECULES OUT OF THE ENDOPHASMIC RETICULUM (ER). SECOND CLEAVAGE IN THE PROPEPTIDE OCCUR IN THE TRANS GOLGI NETWORK (TGN), WHICH IS FOLLOWED BY THE RELEASE OF THE PROPEPTIDE BOUND TO FURIN AND THE ACTIVATION OF FURIN. TISSUE SPECIFICITY: SEEMS TO BE EXPRESSED UBIQUITOUSLY. DOMAIN: CONTAINS A CYTOPLASMIC DOMAIN RESPONSIBLE FOR ITS TGN LOCALIZATION AND RECYCLIKE FROM THE CELL SUFFACE. PTM: THE PROPEPTIDE IS AUTOCATALYTICALLY REMOVED THROUGH AN
respective precursors.
COFACTOR: CALCIUM-DEPENDENT.
ENZYME REGULATION: COULD BE INHIBITED BY THE NOT SECONDLY CLEAVED
```

ö

Gaps

N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
CLEAVAGE (SECOND AUTO-).
CLEAVAGE (FIRST ATTO-).
CELL ATTACHMENT SITE (POTENTIAL).
CELL SURFACE SIGNAL.

553 75 107 500 762

440

CARBOHYD CARBOHYD DISULFID

CARBOHYD

TRANS GOLGI NETWORK SIGNAL.

10C44DD5892EF85D CRC64;

86678 MW;

794 AA;

SEQUENCE

SITE

CHARGE RELAY SYSTEM (BY SIMILARITY). CHARGE RELAY SYSTEM (BY SIMILARITY). CHARGE RELAY SYSTEM (BY SIMILARITY).

POTENTIAL

DOMAIN TRANSMEM ACT\_SITE ACT\_SITE ACT\_SITE DISULFID

CYS-RICH. FURIN.

POTENTIAL

POTENTIAL

PROSITE; PS00136; SUBTILASE\_ASP; 1.
PROSITE; PS00137; SUBTILASE\_HIS; 1.
PROSITE; PS00138; SUBTILASE\_SER; 1.
Hydrolase; Serine protease; Transmembrane; Glycoprotein; Signal;

POTENTIAL

Zymogen; Calcium.

PROPEP CHAIN

MIM: 159530; ...

GO: GO:0004276; F:furin activity; TAS.

GO: GO:0004276; F:furin activity; TAS.

GO: GO:000426; F:furin activity; TAS.

GO: GO:0005809; P:cell-cell signaling; TAS.

GO: GO:0005809; P:proteolysis and peptidolysis; TAS.

GO: GO:0005809; P:proteolysis and peptidolysis; TAS.

GO: GO:0005809; P:proteolysis and peptidolysis; TAS.

InterPro: IPR00281; P:domain.

R InterPro: IPR00209; Peptidase\_S8.

Pfam; PF00183; P:Proprotein; PARTIAL.

PFAM: PF00183; P:Proprotein; PARTIAL.

R PRINTS; PR00717; P:Gomain; I.

SMART; SM00261; FU: Z.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

(ER). IN THE TGN THE SECOND CLEAVAGE IN THE PROPERTIDE COULD LEAD TO THE ACTIVATION OF FURIN.

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY.

INTRAMOLECULAR CLEAVAGE PROBABLY IN THE ENDOPLASMIC RETICULUM

---÷

38888888888888888888888888

PROPEPTIDE.

-

hepatitis hepatitis

089p81 089p83 089p83 089p83 089p83 089p84 088p86 088p82 088p82 088px2 088px2 088px2 088px4 088px4 089px4 08

hepatitis of hepat

080MIS 080MI9 080MJ1 080MJ0 080RI7

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SPTREMBL\_23:\*

Database :

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Scoring table:

Searched:

score:

Perfect

Sequence:

protein

Run on:

STRAIN-K3a;
MEDLINE-95053917; PubMed-7964640;
MEDLINE-95053917; PubMed-7964640;
Yamada N., Manihara K., Mizokami M., Ohba K., Takada A., Tsutsumi M., Genome polyprotein.
Heparitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI\_TaxID=11103; SEQUENCE FROM N.A. STRALM-K3a; Date T.; Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases. Last sequence update) Last annotation update) PRT; 3021 AA 081495 PRELIMINARY; FRI; J 081495; 01.NOV-1996 (TrEMBLrel. 01, Created) 01.NOV-1996 (TrEMBLrel. 01, Last sequ 01.MAR-2003 (TrEMBLrel. 23, Last anno SEQUENCE FROM N.A. RESULT Q81495 DDR READ DR RE

Full length sequence of the genome of hepatitis C virus type 3a:
comparative study with different genotypes.";
J. Gen. Virol. 75:3294-3284(1994).
-1- SUBDUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL: D28917; BAA06044.1; -.
HSSP: P27958; 1HEI.

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

HCV\_capsid. HCV\_core. HCV\_env. HCV\_NSI. DEAD. MEROPS; 529.001; -... MEROPS; U39.001; -... InterPro; IPR001410; D InterPro; IPR002522; H InterPro; IPR002521; InterPro; IPR002519; InterPro; IPR002531;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. sp\_fung1:\*
sp\_human:\*
sp\_invertebrate:\*
sp\_mammal:\* sp\_unclassified:\* sp\_vertebrate:\* sp\_rvirus:\*
sp\_bacteriap:\* sp\_organelle:\* sp\_phage:\* sp\_archea:\* sp\_bacteria:\* sp\_archeap: sp\_rodent: \* sp\_plant:\* sp\_virus:\* sp\_mhc:\*

Description	Q81495 hepatitis Q68307 hepatitis		. —	O8jys3 hepatitis	_	Q8qp87 hepatitis		_	_	Q8qp77 hepatitis	Q8qp84 hepatitis	08qp75 hepatitis
SUMMARIES ID	Q81495 Q68307	Q68308 Q68310	Q8JYS2	Q8JYS3 O8JYR4	Q8QP85	Q8QP87 Q8QP72	Q8QP71	Q8QP74	08000	080P//	Q8QP84	Q8QP75
DB	12	12	12	12	12	12	12	15	7.	7.7	12	12
% Query Match Length DB	3021	4 V V	09	73	100	100	100	100	100	001	100	100
% Query Match	100.0	88.7	88.7	88.7	88.7	88.7	88.7	88.7	. 68	200	68.7	88.7
Score	151	134	134	134	134	134	134	134	134	134	134	134
Result No.	7	m <b>4</b>	ı,	9 ~	<b>60</b> (	10	11	12	57.	# L	CT :	16

```
Gaps
                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VITUS.";
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-: SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
-- LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
-- PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
--- SUBMUT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A SUBMUT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN MAND GLYCOPROTEIN E. THE ENVELOPE CONSISTS OF TWO PROTEINS:
--- PROTEIN A AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL, U337448, AAA65055.1;
--- Interpro; IPR005222; HCV_capsid.
--- Pfam; PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-HCV-BB38;
Songsivilai S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
STRAIN-HCV-C94009;
Sopgivilal S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis
virus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               88.7%; Score 134; DB 12; Length 45; 96.2%; Pred. No. 3.3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           45 AA; 5014 MW; CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 134; DB 12;
Pred. No. 3.3e-13;
0; Mismatches 1;
                                                                                                                                                                                     Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                   Created)
           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; U23746; AAA65053.1; -
InterPro; IPR02522; HCV_caps1d,
Ppfam; PF01543; HCV_caps1d; 1.
Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         88.7%;
ilarity 96.2%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       45 45
45 AA; 5014 MW;
                                     068308;
01-NOV-1996 (TrEMBLrel. 01,
01-NOV-1996 (TrEMBLrel. 01,
01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein.
NON_TER
SEQUENCE 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q68310
Q68310;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   068310
                  SETE THE PROPERTY OF THE PROPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOTER REPRENTED DESTRUCTION OF THE SOLUTION OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE: PS50511; RDRP_VIRAL; 1.
Coat protein; Bryelpe protein; Glycoprotein; Nonstructural protein;
POLyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3021 AA: 328386 MW; A97418FF36C065A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Songsivilal S., Kanistanon D., Kunkitti R.; *Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3021;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   88.7%; Score 134; DB 12; Length 45; 96.2%; Pred. No. 3.3e-13; Live 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0%; Score 151; DB 12;
100.0%; Pred. No. 7.4e-14;
cive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               45 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV 28
LucerPro; IPR002166; HCV_NS5a.
LinterPro; IPR002166; HCV_RGRP.
LinterPro; IPR007095; RNA_pol_DS_PS.
Pfam; PF01543; HCV_caps.d.; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01559; HCV_core; 1.
Pfam; PF01559; HCV_core; 1.
Pfam; PF01538; HCV_LNS1; 1.
Pfam; PF01538; HCV_NS2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PKPQRKTKRNTIRRPQDVKFPGGGQ1 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002522; HCV_capsid, Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                    Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RGRP; 1.
Probom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             45 45
45 AA; 5014 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEXDc; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
nes 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-HCV-BB21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
NON_TER SEQUENCE 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ProDom; I
SMART; SN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
```

968307

Matches

g

ò

SWE

ö

K

ö

Best Loca Matches

ò 8

```
Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-RIG291;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
NON_TER
NON_TER
SEQUENCE 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                             Q8JYR4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8QP85
Q8QP85;
                                                                                                                                                                                                                          Q8JYR4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 8
Q8QP85
                                                                                                                                                                                  RESULT 7
                                                                                                                       g
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-RIG253;
Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/Genbank/DDBJ databases.
-!- SUBMITT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN MAND GILYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/Genbank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN EVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
EMBL, AF506612; AAB3388.1; -
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 134; DB 12; Length 60;
Pred. No. 4.5e-13;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B92CCE7D9C8B9029 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCE7D9C8B90299A CRC64;
                                                                                                                                                                                01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                          60 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CORE PROTEIN.
                    PKPORKTKRNTIRRPODVKFPGGGOI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                               Genome polyprotein (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      60
6725 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88.7%;
96.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
6881 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; AF506613; AAM33389.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity 96,2
nes 25; Conservative
                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     60 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. STRAIN-RIG256;
                                                                                                                                                                                                                                                                    Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 AA;
                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QBJYS3;
01-OCT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                          Q8JYS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QBJYS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loc
Matches
                                                                                                      RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 6
                                                                                                                     ACCOMPANY OF THE PROPERTY OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OBJYS3
ID OE
AC OE
  à
                                       g
                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ó
```

```
ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Magnius L.O.; of hepatitis C virus identified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL, AF506621; ARAM3397.1; -
InterPro, IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   K
                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PETRAIN-602;
MEDLINE-21904745; PubMed-11907242;
Kallinian O., Norder H., Mukomolov S., Magnius L.O.;
Kallinian O., Norder H., Mukomolov S., Magnius L.O.;
An atural intergenotypic recombinant of hepatitis C virus identifi.
In St. Petersburg.";
J. Virol. 76:4034-4043(2002).
I. PURON OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ó
                                                      ö
  61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
  Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6EDC082DAE8CB796 CRC64;
                                                                                                                                                                                                                                                                                                                                          01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JUN-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 134; DB 12;
Pred. No. 5.6e-13;
0; Mismatches 1;
Score 134; DB 12;
Pred. No. 4.6e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                 ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100 AA.
                                                                                                       56
                                                                                                                                   5 PKPQKKTKKNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGQI 26
                                                                                                       1 PKPORKTKRNTIRRPODVKFPGGGVI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGVI
                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                             Genome polyprotein (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 1
73 73 73
73 AA; 8201 MW;
88.78;
96.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88.78;
96.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 96.2
Matches 25; Conservative
                                                      Conservative
                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
```

```
Query Match
                                                                                                                                                                                                                                                                                                                  Q8QP71;
                                                                                                                                                                                                                                                                                                       Q8QP71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        080P74
                                                                                                                                                                                                                                                                              RESULT 11
Q8QP71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 12
    ò
                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                  ö
                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                 Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
"A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg.";
                                                                                                                                                                                                                                                                                                                                                                                                        in St. Petersburg.;
J. Virol. 76:4034-4043(2002)
I. Virol. 76:4034-4043(2002)
I. Sübunt: The Virion of This Virus is a nucleocapsid covered by a Lipoproprein by the Envelope Consists of Two Proteins:
Protein M and Glycoprofin E. The Nucleocapsid is a Complex of Protein C and Mrna (BY SIMILARITY).
                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                        01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  88.7%; Score 134; DB 12; Length 100; 96.2%; Pred. No. 7.8e-13; 1ve 0; Mismatches 1; Indels
                                                                                   Length 100;
                                                100
11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-CCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
                                                                                  Query Match

88.7%; Score 134; DB 12;
Best Local Similarity 96.2%; Pred. No. 7.8e-13;
Matches 25; Conservative 0; Mismatches 1;
                                                                                                                                                                                                                    100 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100 AA.
                                                                                                                                               1 PKPORKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                              STRAIN-233;
MEDLINE-21904745; Pubmed-11907242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=21904745; PubMed=11907242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL, AY070178; AAL58592.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
EMBL; AY070180; AAL58594.1;
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 88.7

9est Local Similarity 96.2

4atches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                    PRELIMINARY;
                                    Polyprotein.
NON_TER 100 1
SEQUENCE 100 AA;
                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein.
NON_TER 10
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-609;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q8QP72
                                                                                                                                                                                                                    080P87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 10
                                                                                                                                                                                            RESULT 9
                                                                                                                                                                                                        080P87
    SO FT SO
                                                                                                                                 à
                                                                                                                                                                                                                                                                   ò
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          080P74;
01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 22, Last squence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                    Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 100;
                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                         100
11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                  01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JUN-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                 88.7%; Score 134; DB 12;
96.2%; Pred. No. 7.8e-13;
11ve 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 134; DB 12;
Pred. No. 7.8e-13;
                                                                                                                                                                                                                                                                                                                                                           100 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0; Mismatches
                                                                                                                                                                                                                                                          1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                     5 PKPORKTKRNTIRRPODVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AYO70194; AAL58608.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88.7%;
96.2%;
                                                                                                                                                                                   Query Match
Best Local Similarity 96.2.
Local Similarity 96.2.
Local 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25; Conservative
                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                     100 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11103;
                                                                                                                                          Polyprotein.
NON_TER 1
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein.
NON_TER
SEQUENCE 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AC DE DIT OC OC OC
```

```
Ralinina O., Norder H., Mukomolov S., Magnius L.O.;

A natural intergenotypic recombinant of hepatitis C virus identified
in St. Petersburg.";
J. Virol. 76:4034-403(2002).

-1 SIBBUIR: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
ILPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPPORTEN E. THE NUCLEOCAPSID IS A COMPLEX OF
EMBL, AV070188; AALSB602.1;
InterPro; IPR002522; HCV_capsid.
Pfam; PF01443; HCV_capsid; 1.
                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               88.7%; Score 134; DB 12; Length 100; 96.2%; Pred. No. 7.8e-13; Live 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100 AA; 11289 MW; F3BC90F5F595F18B CRC64;
              01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JOCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQKKTKKNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                   MEDLINE-21904745; PubMed-11907242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 96.2*
Matches 25; Conservative
                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                          NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
NON_TER 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        080P84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 15
Q8QP84
       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                  STRAIN-631;
MEDLINE-2190745; PubMed-11907242;
Malnina O., Norder H., Mukomolov S., Magnius L.O.;
Ralinina O., Norder H., Mukomolov S., Magnius L.O.;
An natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg.";
J. Virol. 76,4034-4043(2002).
I. SUBONT: THE VIRION OF FHIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL, AV070191; AAL58605.1; ...
InterPro; IPR002522; HCV_capsid.
Pfan; PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kalinina O., Norder H., Mukomolov S., Magnius L.O.;

"A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg "; 7040342002).

J. Virol. 76:4034-403(2002).

-! SUBBUIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA BY SIMILARITY).

EMBL; AX701085; AALS8999.1; -..

InterPro; IPR002222; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88.7%; Score 134; DB 12; Length 100; 96.2%; Pred. No. 7.8e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                    100 100
100 AA; 11289 MW; F3BC90F5F595F18B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
NON_TER 100 100
SEQUENCE 100 AA; 11289 MW; F3BC90F5F595F18B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-CT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               88.7%; Score 134; DB 12;
96.2%; Pred. No. 7.8e-13;
11ve 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PKPORKTKRNTIRRPODVKFPGGGOI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-21904745; PubMed-11907242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 96.2 tes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity 96.2 ies 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
  Hepacivirus.
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein.
NON_TER 1
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        080P80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 13
Q8QP80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TANK REPRESENTANT OF THE PROPERTY OF THE PROPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SO DE RESERVO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

ö

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Kalnina O., Norder H., Mukomolov S., Magnius L.O.;
Ralnina O., Norder H., Mukomolov S., Magnius L.O.;
A natural intergenotypic recombinant of hepatitis C virus identified
in St. Petersburg.";
J. Virol. 76:4034-4043(2002).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
EMBL, ANOTOBI, AND SEASISSSS.1;
-InterPro, IPRO02522; HCV_capsid,
Promission of the complex of the co
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                         01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JUN-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Heparitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match

88.7%; Score 134; DB 12;
Best Local Similarity 96.2%; Pred. No. 7.8e-13;
Matches 25; Conservative 0; Mismatches 1;
100 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-718;
MEDLINE-21904745; Pubmed-11907242;
PRELIMINARY;
```

100 AA.

PRELIMINARY;

**080P77** RESULT 14 080P77 ID 080P7

οg

Search completed: August 7, 2003, 11:19:02 Job time: 26.6364 secs

Page

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 100.0%;
Best Local Similarity 100.0%;
Matches 28; Conservative 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: not relevant TOPOLOGY: not relevant MOLECULE TYPE: protein HYPOTHETICAL: NO ANTI-CENSOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-921-887-30
Sequence 44, Appl
Sequence 130, App
Sequence 189, App
Sequence 189, App
Sequence 191, App
Sequence 192, App
Sequence 192, App
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Appli
Appli
Appli
Appli
Appli
                                                                                                                                             August 7, 2003, 11:07:41; Search time 10.5455 Seconds (without alignments) 112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 190,
Sequence 187,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 175,
Sequence 175,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1
Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 2
Sequence 1
Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence (Sequence (Sequen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-921-887-30
US-08-921-887-23
US-08-290-665A-187
US-08-290-665A-187
US-08-290-665A-190
PCT-US95-10398-187
PCT-US95-10398-190
US-08-865-10398-190
US-08-865-10398-190
US-08-865-10398-190
US-08-921-887-25
US-08-931-887-25
US-08-93-4666-192
US-08-324-977-6
US-08-34-977-6
US-08-34-616-6
US-08-34-616-6
US-08-315-850-6
                                                                                                                                                                                                                                                                                                                                                                                                                                             l number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV 28
                                                                                                                                                                                                                                                                                                                                                                                                  328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                          protein search, using sw model
                                                                                                                                                                                                                                                                                                                                  BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                              US-09-491-146A-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0
                                                                                                                                                                                                                                                                                                                                     Scoring table:
                                                                                                                                                                                                                                                                   Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score
                                                                                                       OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Minimum DB
Maximum DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Database :
                                                                                                                                                                                                                                                                                          Sequence:
                                                                                                                                                    Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
IN AND RESTRICTION ASSISTED LIGATION METHOD FOR MAKING THE SAME
                                                                                                                                         Sequence
Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                        Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
STATE: 6A
COUNTRY: USA
ZIP: 30303-1769
COMPUTER READBLE FORM:
MEDIUW TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN RPELSCATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 151; DB 3;
Pred. No. 1.5e-15;
Mismatches 0;
US-08-083-947-23
US-08-52-037-26
PCT-US93-08638-9
PCT-US93-07088-23
PCT-US95-13660-3
US-08-836-075A-10
US-08-836-075A-14
US-08-836-075A-14
US-08-836-075A-14
US-08-836-075A-14
US-08-836-075A-14
US-08-836-075A-14
US-08-836-075A-14
US-08-84-616-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-244-116B-15
US-08-836-075A-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-315-850-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sequence 30, Application US/08921887
FERRAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN ANI
TITLE OF INVENTION: ENDONUCLEASE ASSINUMBER OF SEQUENCES:
NUMBER OF SEQUENCES:
ADDRESSEE: JONESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           03063-0380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
CLASSIFICATION: 435
ATORNEZ/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063
TELECOMMUNICATION INFORMATION:
TELEFAX: 404-818-3709
INFORMATION FOR SED ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acids
TYPE: amino acids
TYPE: amino acids
TYPE: amino acids
```

ö

Gaps

ö

Length 28; Indels

à

```
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEPATITIS C VIRUS AND THE USE OF REAGENTS DERIVED FROM THESE SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         R.H.
NUCLEOTIDE AND DEDUCED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                                                                                                      SUFTWARE: WORDERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT THE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   88.7%; Score 134; 96.2%; Pred. No. 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MILLER, R.H. AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: US/08/290,665A FILING DATE: 15-AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 188, Application US/08290665A Patent No. 5882852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PC-DOS/MS-DOS
                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS
                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
                                                                                                                                                                                                                                                                                                                                                                                                        TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 187:
                                                                                                                                                                                                                                                                                                                                                       REFERENCE/DOCKET NUMBER: 20
                                                                                                                                                                                                                                                                                                                                                                                           TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDIUM TYPE: FLOPPY DISK COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS SOFTWARE: WORDPERFECT 5.1
                                   E: MORGAN & FIN
345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 191 amino acids
TYPE: amino acids
STRANDEDNESS: unknown
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ouery Match
Best Local Similarity 96.2
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: MORGAN CORPERT: 345 PARK AVENUE
                CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INDIVIDUAL ISOLATE: HR
US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
APPLICANT: BUKH, J., MIL
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCL
TITLE OF INVENTION: CORE
TITLE OF INVENTION: CORE
TITLE OF INVENTION: AND
TITLE OF INVENTION: SEQUENCES OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION DATA
NUMBER OF SEQUENCES;
                                                                                     NEW YORK
: USA
                                                                      NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGINAL SOURCE
                                                                                                                   10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 4
US-08-290-665A-188
                                                               CITY: NEW
STATE: NE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ά
                                                                                                                                                            APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEPATITIS C VIRUS AND THE USE OF REAGENTS DERIVED FROM THESE SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 28;
                                                                                                                                                                                                                                                                                                                                                                            COMPUTER REACABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATION SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                           ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 134; DB 3;
Pred. No. 4.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. 4.5e
0; Mismatches
           Sequence 187, Application US/08290665A
Patent No. 5882852
GENERAL INFORMATION:
APPLICANT: BUKH, J, MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 23, Application US/08921887
Patent No. 6030771
GENERAL INFORMATION;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 0306
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 96.2%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO.
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein HYDOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                 30303-1769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORIGINAL SOURCE
                                                                                                                                                                                                                                                                                                                                     S
                                                                                         US-08-921-887-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18-921-887-23
                  a
```

ò qq

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                      AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEPATITIS C VIRUS AND THE USE OF REAGENTS DERIVED FROM THESE SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 5; Length 191;
                                                                                       Length 191;
                                                                                                                                     1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            88.7%; Score 134; DB 5; 96.2%; Pred. No. 3.6e-12;
                                                                                     Score 134; DB 2;
Pred. No. 3.6e-12;
0; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                    R.H.
NUCLEOTIDE AND DEDUCED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      UMBER: PCT/US95/10398
15-AUG-1995
                                                                                                                                                                                                                                                                                                             PCT-US95-10398-187
; Sequence 187, Application PC/TUS9510398
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CURRENT APPLICATION DATE:

CURRENT APPLICATION DATE:

FILING DATE:

CIACT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE: 15-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 187:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-
SOFTWARE: WORDPERFECT 5.1
                                                                                     Query Match
Best Local Similarity 96.2%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: homosaplens;
INDIVIDUAL ISOLATE: HK10
PCT-US95-10398-187
  ) ORGANISM: homosapiens
) INDIVIDUAL ISOLATE: DK12
US-08-290-665A-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: MUCITILE OF INVENTION: AMINITILE OF INVENTION: CORRITTLE OF INVENTION: SEQUENCES: 263 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: RICHARD W. BORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET: 345 PARK
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10154
                                                                                                                                                                                                                                                                                          RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEAPITILE C VIRUS NOT THE USE OF REAGENYE DERIVED FROM THESE SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .
                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 2; Length 191; 3.6e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 190, Application US/08290665A; Patent No. 5882852; GENERAL INFORMATION:
APPLICANT: BURKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF TITLE OF INVENTION: AND THE USE OF ISOLATES OF
TITLE OF INVENTION: AND THE USE OF REAGENTS DE
TITLE OF INVENTION: AND THE USE OF REAGENTS DE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC ME
                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 88.7%; Score 134; Best Local Similarity 96.2%; Pred. No. 3 Matches 25; Conservative 0; Mismatch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTONNEY,AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/POCKET NUMBER: 36,459
REFERENCE/POCKET NUMBER: 36,459
REFERENCE/POCKET NUMBER: 3026-4116
TELECHONE: (212) 758-4800
ATTORNEY AGENT INFORMATION:
NAME: RICHARD W BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELECHONE: (212) 758-480
TELERA: (212) 751-6849
TELERA: (212) 751-6849
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEX: 421792
INFORMATION FOR SEO ID NO: 190:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                      ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S52
                                                                                                                                                                                                                                                                                          SS: unknown unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NEW YORK
: NEW YORK
RY: USA
                                                                                                                                                                                                                                                                                                               TOPOLOGY: unk
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 5
US-08-290-665A-190
                                                                                                                                                                                                                                                                                                                                                                                                    US-08-290-665A-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
```

```
Sequence 44, Application US/08836075A
Patent No. 6180768
GENERAL INFORMATION:
APPLICANT: MARRIENS, GEERT
APPLICANT: STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: AGENTS
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
AND ADDRESS:
  AND THE USE OF REAGENTS DERIVED FROM THESE SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88.7%; Score 134; DB 5; thatity 96.2%; Pred. No. 3.6e-12; Conservative 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              UMBER: PCT/US95/10398
15-AUG-1995
                                                                                                                                                                                                                                                                                                                                  APLICATION NOMBER: PC17(0855/1035)
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-411
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5: ARNOLD, WHITE & DURKEE P.O. BOX 4433
                                                                                                                                                                                                                                      COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                               MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 42.192
INFORMATION FOR SEQ ID NO: 190:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                         SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: homosapiens
: INDIVIDUAL ISOLATE: DK12
PCT-US95-10398-190
                                                                             ADDRESSEE: MORGAN & FIN
STREET: 345 PARK AVENUE
CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANTER
ORGANTER
ORGANTER
                                     NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 25; Conserva
                                                                                                                                        NEW YORK
: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        77210-4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STREET: P.O. BOOTTY: HOUSTON STATE: TEXAS
                                                                                                                                                                           10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-836-075A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                        TITLE OF INVENTION: WUCLECTIDE AND DEDUCED
TITLE OF INVENTION: ORE GENES OF ISOLATES OF HEBATITIES C VIRUS
TITLE OF INVENTION: ORE GENES OF ISOLATES OF HEBATITIES C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R.H.
NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  88.7%; Score 134; DB 5; Length 191; larity 96.2%; Pred, No. 3.6e-12; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Mismatches
                                                                                                        APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      омыЕR: PCT/US95/10398
15-AUG-1995
                                                                  Sequence 188, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/1039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REFERENCE/DOCKET NUMBER: 2026-411
TELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 190, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEX: 421792
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                       STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SS: unknown unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Thes 25; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: amino acid
STRANDEDNESS: un)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE: 15
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGINAL SOURCE:
                                              PCT-US95-10398-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PCT-US95-10398-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PCT-US95-10398-188
                                                                                                                                                                                                                                                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

ô

Gaps

```
NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEBATITIES C VIRUS
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88.7%; Score 134; DB 4; Length 319; 96.2%; Pred. No. 6.4e-12; Live 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-00-290-665A-189
Sequence 189, Application US/08290665A
Batent No. 5882852
GENERAL INFORMATION:
APPLICANT: BURH, J., MILLER, R.H. AND
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCE
TITLE OF INVENTION: AMINO ACID SEQUENCES
TITLE OF INVENTION: CORE GENES OF ISOLATION: TITLE OF INVENTION: SEQUENCES IN DIAGNOS;
TITLE OF INVENTION: SEQUENCES IN DIAGNOS;
CORRESSEDENCES: 263
CORRESSEDENCES: MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; LOCATION: (233)..(233)
; OTHER INFORMATION: Xaa is any amino acid
US-08-635-886C-230
                                                                                                                                                                                              LOCATION: (171)..(172)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                        LOCATION: (177)..(177)
OTHER INFORMATION: Xaa is any amino acid
     ION: (161)..(161)
INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                           LOCATION: (174)..(174)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LUCCATION: (232)..(232)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
                                                                                                    LOCATION: (167)..(167)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15 AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 20,
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : (212) 758-4800
(212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E: MORGAN & FIN
345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 96.2
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                           NAME/KEY: MISC_FEATURE
                                                                                                                                                                     NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                  NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                               NAME/KEY: MISC_FEATURE LOCATION: (177)..(177)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEY: MISC_FEATURE LOCATION: (233)..(233)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NEW YORK

NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
APPLICANT: LEROUX ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: DELEYS, Robert
APPLICANT: DELEYS, Robert
TITLE OF INVENTION: URBUSOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VIRUS
FILE REFERENCE: 2752-18
FILE REFERENCE: 2752-18
FILE REPERENCE: 2752-11-04
FRIOR FILING DATE: 1994-10-28
FRIOR FILING DATE: 1991-11-04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88.7%; Score 134; DB 3; Length 319; 96.2%; Pred. No. 6.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Indels
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Mismatches
                                                                                     PRICE APPLICATION NUMBER: US/08/836,075A
FILING DATE: 21 Apr 1997
PRIOR APPLICATION DATA: 1997
PRIOR APPLICATION NUMBER: PCT/EP95/04155
FILING DATE: 23 Oct 1995
PRIOR APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
PRIOR APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY AGENT INPORMATION:
NAME: KANMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/POCKET NUMBER: 29,775
REFERENCE/POCKET NUMBER: 29,775
REFERENCE/POCKET NUMBER: 1NNS:004
INPORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LOCATION: (156)..(157)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPORKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 230, Application US/08635886C Patent No. 6555114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LOCATION: (144)..(144)
OTHER INFORMATION: Xaa is any
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQ ID NOS: 286
SOFTWARE: PatentIn version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : 319 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 88.7
Best Local Similarity 96.2
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; MOLECULE TYPE: peptide US-08-836-075A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 10
US-08-635-886C-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQ ID NO 230
LENGTH: 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

ô

```
GENERAL INFORMATION:
APPLICANT: KUDYAROV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: BUDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAM
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
                                                                          Gaps
                                                                          ö
                                Score 133; DB 5; Length 191;
Pred. No. 5.1e-12;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 85.4%; Score 129; DB 3; Best Local Similarity 92.3%; Pred. No. 2.4e-12; Matches 24; Conservative 0; Mismatches 2
                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                             5 PKPQRKTKRNTIRRPQDIKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQKKTKRNTYRRPQDVKFPGGGQI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 191, Application US/08635886C Patent No. 655514 GENERAL INFORMATION: APPLICANT: LEROUX-ROBLS, Geert APPLICANT: DELEYS, ROBERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPALIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                         Sequence 25, Application US/08921887 Patent No. 6030771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AVII-SENSE: not relevant MOLECULE TYPE: protein HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RECISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-918-3700
                          Query Match

Best Local Similarity 92.3%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS: LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              404-818-3799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FRAGMENT TYPE: 1
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 14
US-08-635-886C-191
                                                                                                                                                                                                                     US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUCLEOTIDE AND DEDUCED
ANINO ACID SEQUENCES OF THE ENVELOPE 1 AND
COPE GENES OF ISOLATES OF HEPATITIES C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                       Length 191;
                                                                                                                                                                                                                                 88.1%; Score 133; DB 2;
92.3%; Pred. No. 5.1e-12;
tive 1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CURRENT APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: PCI/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                  Sequence 189, Application PC/TUS9510398
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026
TELECOMMUNICATION INFORMATION:
TELEPAN: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 189:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  E: FLOPPY DISK
IBM PC COMPATIBLE
INFORMATION FOR SEQ ID NO: 18: SEQUENCE CHARACTERISTICS: LENGTH: 191 amino acids TYPE: amino acid STRANDEDNESS: unknown TOPOLOGY: unknown ORIGINAL SOURCE:
                                                                                                                                                                                                             Query Match
Best Local Similarity 92.30,
...Ahes 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : 345 PARK AVENUE
NEW YORK
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH: 191 amino acids
                                                                                                                                                     ORGANISM: homosaplens
; INDIVIDUAL ISOLATE: S2
US-08-290-665A-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: N
TITLE OF INVENTION: A
TITLE OF INVENTION: C
TITLE OF INVENTION: A
TITLE OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: amino acid
STRANDEDNESS: un)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                         PCT-US95-10398-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PCT-US95-10398-189
                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                              Q
```

us-09-491-146a-30.rai

```
Sequence 192, Application US/08635886C

Sequence 192, Application US/08635886C

Patent No. 655114

GENERAL INFORMATION:

APPLICANT: LEBEOUX-NOELS, Geert

APPLICANT: DELEYS, Robert

TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C

TITLE OF INVENTION: VIRUS

TITLE OF INVENTION: VIRUS

TITLE OF INVENTION: VIRUS

CURRENT APPLICATION NUMBER: US/08/635,886C

CURRENT APPLICATION NUMBER: D196-04-25

PRIOR FILING DATE: 1994-10-28

PRIOR FILING DATE: 1994-10-28

PRIOR FILING DATE: 1993-11-04

NUMBER OF SEQ ID NOS: 286

SEQ ID NO 192

LENGTH: 450

TYPE: PRIOR FILING PARENTIN OF SEQ ID NO 152

TYPE: PRIOR FILING DATE: PRIOR FILING PRIOR FILING DATE: PRIOR FIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: UMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VINUS
FILE REFERENCE: 275-18
CURRENT FILING DATE: 1996-04-25
PRIOR APPLICATION NUMBER: DFT/EP94/03555
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1994-10-28
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PALENTI VETSION 3.1
SEQ ID NO 191
LENGTH: 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 129; DB 4; Length 450;
Pred. No. 5e-11;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 85.4%; Score 129; DB 4; Length 450; Best Local Similarity 92.3%; Pred. No. 5e-11; Matches 24; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 85.4%;
Best Local Similarity 92.3%;
atches 24; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: hepatitis C virus US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

qq

Search completed: August 7, 2003, 11:23:52 Job time: 10.6364 secs

```
Sequence 44, Appl
Sequence 1, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 10, Applance 10, Applance 11, Applance 12, Applance 14, Applance 14, Applance 16, Applance 152, Applance 153, Applance 153, Applance 153, Applance 153, Applance 154, Applance 155, Applance 154, Applance 155, Applance 154, Applance 155, Applance 154, Applance 155, Applance 
                                                                                                                            August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0 US-09-851-138-44
0 US-09-851-138-10
0 US-09-758-308-1
0 US-09-756-875-8
1 US-09-751-397-77
0 US-09-921-397-77
0 US-09-921-397-78
0 US-09-921-397-78
0 US-09-921-397-78
0 US-09-951-138-46
0 US-09-951-138-46
0 US-09-899-046-152
0 US-09-899-046-152
0 US-09-899-046-152
0 US-09-899-046-142
                                                                                                                                                                                                                                                                                                                                                                                                          tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                     151
1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV 28
                                                                                                                                                                                                                                                                                                                                                                  451899 seqs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                      OM protein – protein search, using sw model
                                                                                                                                                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                 US-09-491-146A-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     319
444
744
1003
1103
1113
1136
1166
1169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
                                                                                                                                                                                                               Title:
Perfect score:
                                                                                                                                                                                                                                                                                                     Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Database :
                                                                                                                                                                                                                                                                                                                                                                    Searched:
                                                                                                                                                                                                                                                         Sequence:
                                                                                                                              Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Result
```

```
APPLICANT: MAERIENS, GEERT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Appli
e 1, Appli
20, Apr
                                                     2, Appli
36, Appli
112, Appl
112, Appl
112, Appl
52, Appl
52, Appl
52, Appl
52, Appl
52, Appl
50, Appl
60, Appl
60, Appl
70, Appl
                                                                                                                       Sequence 12, Sequence 18, Sequence 50, Sequence 52, Sequence 54, Sequence 54, Sequence 54, Sequence 56, Sequence 50, Sequence 23, Sequence 20, No. Sequence 23, S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 4, Apr
                                                                              Sequence Sequence 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1
Sequence 2
Sequence 3
Sequence 3
Sequence 1
Sequence 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
US-09-878-281-44
US-09-1929-955-2
US-09-104-966-2
US-09-194-949-12
US-09-191-138-76
US-09-851-138-76
US-09-851-138-76
US-09-899-046-54
US-09-899-046-54
US-09-899-046-54
US-09-899-046-144
US-09-899-046-144
US-09-878-281-52
US-09-878-281-54
US-09-878-281-54
US-09-878-281-54
US-09-878-281-54
US-09-891-303-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-952-572-9
US-09-929-955-1
US-09-747-419-20
US-09-891-894-3
US-10-104-966-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-10-259-275-20
US-09-238-076-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CORRESPONDENCE ADDRESS:
RADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 44, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATTORNEY/AGENT INFORMATION
       NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                     2894
2894
2985
3011
3011
3011
3011
3011
    US-09-851-138-44
       RESULT 1
```

INFORMATION FOR SEQ ID NO:

```
Sequence 1, Application US/09758308
Patent No. US20020090607a1
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: HOWARD N. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C V
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 14114.034902
CURRENT APPLICATION NUMBER: US/09/758,308
CURRENT FILING DATE: 2001-01-10
PRIOR FILING DATE: 1999-07-10
NUMBER OF SEQ ID NOS: 5
SOFWARE: PATENTIN VEFSION 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: MICTOSOFT WORD 6.0 / ASCII text output CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,075
FILING DATE: cUNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    84.1%; Score 127; DB 10;
1larity 92.3%; Pred. No. 3.8e-11;
Conservative 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 127; DB 9;
Pred. No. 4.8e-11;
                                                                                    ADDRESSE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COMPRRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: FLORPY disk
COMPUTER: IBM PC COMPATIBLE
                                                                                               WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 74 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ch 84.1%;
l Similarity 92.3%;
24; Conservative
                                                   CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; ORGANISM: Hepatitis C Virus US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
Matches 24; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: JOLIVET, MICHEL
APPLICANT: DALINE, FRANCOIS
APPLICANT: DALINE, FRANCOIS
APPLICANT: DALINE, LAURENT
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGEN
TITLE OF INVENTION: TREATING AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
TITLE OF INVENTION: TREATING AND
FILE REFERENCE: 103959
CURRENT APPLICATION NUMBER: US/09/389,756
PRIOR FILING DATE: 1999-09-07.
PRIOR PAPLICATION NUMBER: BALLIER APPLICATION NUMBER: CFT/FR98/00442
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DETECTING, PREVENTING AND
                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                               Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
TITLE: of Hepatitis Virus
JOURNAL: Proc. NATI. Acad. Sci. U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match

84.1%; Score 127; DB 15; Length 44;
Best Local Similarity 92.3%; Pred. No. 2.2e-11;
Matches 24; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                             1; Indels
                                                                                                                                                                                                                                                                                Query Match

88.7%; Score 134; DB 10;
Best Local Similarity 96.2%; Pred. No. 1.9e-11;
Matches 25; Conservative 0; Mismatches 1;
SEQUENCE CHARACTERISTICS:
LENGTH: 319 annho acids
LENGTH: 319 annho acids
TYPE: annho acid
TYPE: annho acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-851-138-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4 PKPQRKTKRNTNRRPQDVKFPGGGQI 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 10, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MARRTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : Sequence 1, Application US/10367677 'ublication No. US20030118604A1 ENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; RELEVANT RESIDUES: 2 TO 45 US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PAGES: 3392-3396
DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
STUTYER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 108;
                                                               Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER: IBM PC compatible
COMPATION SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 127; DB 10;
Pred, No. 5.5e-11;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-851-138-14
                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                1 PKPORKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                             Sequence 14, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; Sequence 78, Application US/09921397; Patent No. US20020111484A1; GENERAL INFORMATION; APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                               APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET: P.O. BOX 4433
                                                             Query Match
Best Local Similarity 92.3%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS
; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 24; Conserva
                                                                                                                                                                                                                                                       RESULT 7
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 8
                                                                                                                                                                                          g
                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polypeptides selected from a
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 97;
                                                                            Sequence 8, Application US/09756875;
Patent No. US20020150990A1;
GENERAL INFORMATION:
APPLICANT: PIKE, IAN
TILLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sulte 701-E Columbia Square
STREET: 555 13th Street, N. W.
                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 127; DB 10;
Pred. No. 5.2e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 77, Application US/09921397
Sequence 77, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
TITLE OF INVENTION: pathogenic strain of the
TITLE OF INVENTION: pathogenic strain of the
TITLE OF INVENTION: applications thereof
FILE REFERENCE: B4809A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
FRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 77
TYPE: PRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-40G-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 1808-157A
REFERENCE/CDOCKET NUMBER: 1808-157A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPOTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 92.3%;
Matches 24; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; MOLECULE TYPE: peptide US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                                                                                                                                                                                                                                                                                               STATE: D. C. COUNTRY: U. S. ZIP: 20004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-921-397-77
                                                               US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        엄
```

```
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: B94870166.9
FILING DATE: CURROWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 138;
                              Indels
                       2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 127; DB 10;
Pred. No. 7.6e-11;
0; Mismatches 2;
                    0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CORRESPONDENCE ADDRESS:
RADRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-851-138-60
                                                                                                                  5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PKPORKTKRNTNRRPODVKFPGGGGI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 152, Application US/09899046
Publication No. US20030008274A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                       Sequence 60, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 138 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                          LIEVEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 92.3%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid
                24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-899-046-152
                                                                                                                                                                                                                                       RESULT 10
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT
                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and TITLE OF INVENTION: applications thereof
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES: 201

CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUGTON
STATE: TEXAS
COUNTRY: USA
IP: 77210-4433
COMPUTER READABLE FORM:
WEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
APPLICATION DATA:
APPLICATION DATA: 09-M8T: 09-M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 127; DB 10;
Pred. No. 7.5e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.le-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match

84.1%; Score 127; DE
Best Local Similarity 92.3%; Pred. No. 6.1e
Matches 24; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: 08/836,075
FILING DATE: CURNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                   FILE REFERENCE: B4809A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
FRIOR RAPLICATION NUMBER: EP 00402225.7
FRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SEQ ID NO 78
SEQ ID NO 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATTORNEY/AGENT INFORMATION:
NAME: KAMMEREN, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PKPORKTKRNTNRRPODVKFPGGGOI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 46, Application US/09851138
Publication No. US20020183508Al
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 137 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                        ; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: MAERTENS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

Gaps

ó

Indels

Score 127; DB 11; Pred. No. 9.5e-11; 0; Mismatches

84.1%;

```
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTKRNTIRRPODVKFPGGGVI 26
                                                                                              APPLICATION NUMBER: US/09/899,046
FILING DATE:
FRIOR PAPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                       169 amino acids
                                                                                                                                                                         FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
                                                                           CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                         24; Conservative
                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein
                                                                                                                                                                                                                                                      amino acid
                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY:
                                                                                                                                                                                                                                                                                                              US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                           Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 166;
                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                           84.1%; Score 127; DB 11;
92.3%; Pred. No. 9.3e-11;
Live 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 127; DB 11;
Pred. No. 9.3e-11;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PKPQRKTKRNINRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                     CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 152, Application US/09878281; Publication No. US20030032005A1
GENERAL INFORMATION: APPLICANT: APPLICANT:
                                                                                              FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 42, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           84.1%;
92.3%;
                                                                                                                                                                                                           LENGTH: 166 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 92.3*
                                                                                                                                                    FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 92.3
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                              TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            , TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT:
TITLE OF INVENTION: NG
TITLE OF INVENTION: 96
NUMBER OF SEQUENCES: 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                        US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
```

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy,
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIOM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New sequences of hepatitis C virus
genotypes for diagnosis, prophylaxis and therapy.
270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 127; DB 11;
Pred. No. 9.5e-11;
0; Mismatches 2;
1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PKPORKTKRNTNRRPODVKFPGGGOI 30
                                                                                                                             Sequence 44, Application US/09899046
Publication No. US20030008274A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO: 44:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 92.3%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-878-281-42
Sequence 42, Application US/098
Publication No. US20030032005A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      169 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT:
TITLE OF INVENTION: N
TITLE OF INVENTION: 9
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
```

```
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
COMPUTER: IBM PC Compatible
COMPUTER: IBM PC Compatible
CORRATING SYSTEM: PC-DOS/MS-DOS
CORRATING SYSTEM: PC-DOS/MS-DOS
CORRATION DATA:
PRICE DATE:
PRICE APPLICATION DATA:
APPLICATION TOWART TO
```

us-09-491-146a-29.rpr

genome polyprotein polyprotein polypeptide - hepa genome polyprotein genome polyprotein

JO0881 PC2219 PC2219 GNWYTC GNWYCH S18030 JC5620 JC5620 JC1303 GNWYJ8 S41343 S41352 S41352 S41352 S41353

874 3010 3010 3010 3011 3011 115 115 115 115 105 105

138 138 138 138 138 133 137 137 133

ALIGNMENTS

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR\_76:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\*
4: pir4:\*

Minimum DB seq length: 0 Maximum DB seq length: 2000000000 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

rip	genome polyprotein	genome polyprotein
SUMMARIES	GNWVC3 841353 841353 841354 841354 841341 841341 841342 841344 841350 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 841344 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84134 84	200883
DBB		7
gth	3011 1008 1008 1112 1114 1115 1115 1115 1115 1115 1115	874
% Query Match	W & & & & & & & & & & & & & & & & & & &	6
Score		138
Result No.		29

1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28

ò

T. T. L.	an	R;Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C. Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991 A;Title: Genetic organization and diversity of the hepatitis C virus. A;Reference number: A39166; MUID:91172826; PMID:1848704 A;Accession: A39166 A;Accession: A30166 A;Residues: 1.3011 <cho></cho>	A;Cross-references: GB:M62321; NID:g329873; PIDN:AAA45676.1; PID:g329874 R;Chan, S,W.; McOmlsh, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yaj J. Gen. Virol. 73, 1131-1141, 1922 A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationshij A;Reference number: PQ0393; MUID:92268871; PMID:1316939 A;Accession: PQ04033; MUID:92268871; PMID:1316939	A; Residues: 1577-1633 CRRA. A; Cross-references: DDBJ: D10128 A; Experimental source: 1solates E-b16 A; Accession: P00404 A; Status: preliminary A; Molecule type: genomic RNA A; Residues: 1577-1633 CRRA. A; Residues: 1577-1633 CRRA. A; Experimental source: 1solates E-b17	C;Superfamily: hepatitis C virus genome polyprotein C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstru F;115/Product: capsid protein C #status predicted <epn> F;120-389/Product: envelope protein M #status predicted <epn> F;320-729/Product: major envelope protein NSI #status predicted <nsi>F;330-729/Product: nonstructural protein NSI #status predicted <nsi>F;730-1006/Product: nonstructural protein NSI #status predicted <ns2> F;1210-1237/Region: nucleotide-binding motif A (P-loop) F;1312-1317/Region: nucleotide-binding motif B F;1316-1319/Region: DEXH motif</ns2></nsi></nsi></epn></epn>	F;1616-1862/Product: nonstructural protein NS4a #status predicted <n4a> F;1863-2013-Yproduct: nonstructural protein NS4b #status predicted <n4b> F;2014-3011/Product: nonstructural protein NS5 #status predicted <ns5> F;2014-3011/Product: nonstructural protein NS5 #status predicted <ns5> F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2( Ouery Match 95.5%; Score 147; DB 1; Length 3011; Best Local Similarity 92.9%; Pred. No. 3.6e-12; Matches 26; Conservative 2; Mismatches 0; Indels 0; Gaps 0;</ns5></ns5></n4b></n4a>
----------------------------------------------	----	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

```
Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                             A.Molecule type: genomic RNA
A.Residues: 1-108 <VAN>
A.Cross-references: BrBL:229460
A.Cross-references: genotype 2, N6
C.Superfamental source: genotype 2, N6
C.Superfamily: hepatitis C virus genome polyprotein
C.Keywords: capsid protein; core protein; polyprotein
F:1-108/Product: core protein *status predicted <MAI>
                                                                 A; Reference number: S41341
A; Accession: S41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Accession: S41348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S41348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (genotype 2, N2) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N2
C;Bate: 19*May-1994 sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41353
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LipA.
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 2, N4
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: $41355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N.Contains: core protein
C;Species: hepatitis C virus
A;Variaty; genorype 2, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41357
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994 A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by Lipa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A.Cross-references: EMBL:229456
A.Experimental source: genotype 2, N2
Superfamily: hepatitis C virus genome polyprotein
Sywords: capaid protein; core protein; polyprotein
-108/Product: core protein #status predicted <a href="Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:Mailto:M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Yeerimental source: genotype 2, N4
.pperfamily: hepatitis C virus genome polyprotein
., Acywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <AMT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 138; DB 2;
Pred. No. 2.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 138; DB 2;
Pred. No. 2.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3; Mismatches
        32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQKRNQRNTNRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PKPQKKNKRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
Best Local Similarity 85.7%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A, Accession: S41355
A, Molecule type: genomic RNA
Residues: 1-108 <VAN>
A, Cross-references: EMBL: 229458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 85.7
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: ger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
```

```
N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, N6
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 541348
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
aubmitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by Lippa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: qenotype 5, N5
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S4131,
E;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5, N5) (fragment)
        Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A. Cross-references: EMBL:229451
A. Experimental source: genotype 1, N6
C. Superfamily: hepatitis C virus genome polyprotein
C. Steywords: capsid protein; core protein; polyprotein
F. 1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A: Experimental source: genotype 5, N5
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-112/Product: core protein #status predicted <MAT>
     Score 138; DB 2;
Pred. No. 2.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 138; DB 2;
Pred. No. 2.4e-12;
3; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus (genotype
                                 Pred. No. 2.46
3; Mismatches
                                                                                                                                         1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
Query Match
Best Local Similarity 85.7%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        89.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 85.7%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Cross-references: EMBL: 229474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: genomic RNA A; Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: genomic RNA A; Residues: 1-112 <VAN>
```

ö

Length 112;

Score 138; DB 2; Pred. No. 2.5e-12;

m

ö

Gaps

ö

Indels

Length 114;

Matches

a

à

```
C.Species: hepatitis C virus
A; Variety: genotype 1, N10
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41342
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A; Reference number: S41341
A; Reference number: S41341
A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Variety; genotype 5, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41368
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41341
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000 C;Accession: S41369
R;Aan Boorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
A;Accession: S41360
A;Experimental source: genotype 5, N3
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <MATP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Residues: 1-114 <VAN>
A; Residues: 1-114 <VAN>
A; Residues: 1-114 <VAN>
A; Cross-references: EMBL: Z29471; NID:g443904; PIDN:CAA82609.1; PID:g443905
A; Experimental source: genotype 5, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <NAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
A; Cross references: EMBL: 229445; NID:g443852; PIDN:CAA82583.1; PID:g443853
A; Experimental source: genotype 1, N10
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               polyprotein - hepatitis C virus (genotype 1, N10) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus (genotype 5, N2) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 2;
Pred. No. 2.5e-12;
3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 138; DB 2;
Pred. No. 2.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 85.7%;
Matches 24; Conservative 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           89.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N;Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Accession: S41368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S41370
genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
N;Contains: core protein
N;Contains: core protein
N;Contains: core protein
S, Nariety: genotype 5, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41370
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the RMBL Data Library, January 1994
.'Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
eference number: S41370
....
                                                                                                                                                                                          RESULT 7
S41341
S401341
Senome polyprotein - hepatitis C virus (genotype 1, N1) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N1
C;Species: hepatitis C virus
A;Variety: genotype 1, N1
C;Accession: S41341
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
A;Bedrence on hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference on hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Molecule type: genomic RNA
A;Residues: 1-114 <VAN>
A;Residues: 1-114 <VAN>
A;Cross-references: EMBL:229473; NID:g443908; PIDN:CAA82611.1; PID:g443909
A;Experimental source: genotype 5, N4
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A,Cross-references: EMBL:229444, NID:9443850; PIDN:CAA82582.1; PID:9443851 C,Experimental source: genotypel, NI C;Superimental source: genotypel, NI C;Superimently: hepatitis C virus genome polyprotein C;Keywords: capsid protein; core protein; polyprotein F;1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5, N3) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
        Indels
     1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 138; DB 2;
Pred. No. 2.5e-12;
3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 138; DB 2;
Pred. No. 2.5e-12;
3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S41369
genome polyprotein - hepatitis C virus (genotype N:Contains: core protein C;Species: hepatitis C virus A;Variety: genotype 5, N3
        Mismatches
                                                                                        1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                            1 PKPOKRNORNINRRPQDVKFPGGGQIVG
        3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     89.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 89.6%;
Best Local Similarity 85.7%;
Matches 24; Conservative
     24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          lolecule type: genomic RNA
Residues: 1-112 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Accession: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Simil
Matches 24; (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S
```

g ð

ö

Gaps

ö

Indels

RESULT 9

g

à

Length 114;

```
yecume profession imparating vitus (genouppe 4, ns) (itayment)
yecume profession imparating vitus
A; Variety igenotype 2, N3
A; Variety; genotype 2, N3
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 541354
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: 54131
A; Accession: 841354
A; Accession: 841354
A; Molecule type: genomic RNA
A; Residues: 1-115 < VAN>
A; Cross-references: EMBL: 229457
A; Experimental source: genotype 2, N3
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; #status predicted < MAT>
F; 1-115/Product: core protein #status predicted

    hepatitis C virus (genotype 2, N3) (fragment)

                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local
Matches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Search co
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                               N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, N8
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Coession: $41350
an Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
mitted to the EMBL Data Library, January 1994
A; Reference number: $41350
A; Accession: $41350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A)Cross-references: EMBL:229447; NID:9443856; PIDN:CAA82585.1; PID:9443857 A)Experimental source: genotype 1, N2 C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F;1-115/Product: core protein #status predicted <WAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A Molecule type: genomic RNA
A; Residues: 1-115 < VAN>
A; Residues: 1-115 < VAN>
A; Cross-references: EMBL: 229453; NID: 9443868; PIDN: CAA82591.1; PID: 9443869
A; Experimental source: genotype 1, N8
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted < MAT>
                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (genotype 1, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (genotype 1, N8) (fragment)
                                               Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89.6%; Score 138; DB 2; Length 115;
llarity 85.7%; Pred. No. 2.5e-12;
Conservative 3; Mismatches 1; Indels
                                                                                       1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
F;1-115/Product: core protein #status predicted <MAT>
                                            Score 138; DB 2;
Pred. No. 2.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 138; DB 2;
Pred. No. 2.5e-12;
3; Mismatches 1
                                                                                       Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. 2.5e
3; Mismatches
                                                                                                                             1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                        1 PKPOKRNORNTNRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             89.68;
                                            89.6%;
85.7%;
                                      Query Match 89.6
Best Local Similarity 85.7
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: genomic RNA A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 24; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 24; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Accession: S41344
                                                                                                                             ò
                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

RESULT 14 S41354

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (genotype 1, N3) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N3
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41345
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
B;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
A;Accession: S41345
A;Molecule type: genomic RNA
A;Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Cross-references: EMBL:229448; NID:9443858; PIDN:CAA82586.1; PID:9443859
A;Experimental source: genotype 1, N3
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; corre protein; polyprotein
F;1-115/Product: core protein #status predicted <MAT>
                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                            Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 115;
Score 138; DB 2; Length 11
Pred. No. 2.5e-12;
3; Mismatches 1; Indels
                                                                                                                 1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7, 2003, 11:21:48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 Similarity 85.7%;
24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           89.68;
                                           Best Local Similarity 85.7
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     August
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   completed: Augus:
le : 9.54545 secs
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

US-09-491-146A-29 154 1 PKPQKRNORNTNRRPQDVKFPGGGQIVG 28

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched:

127863 . number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

		Description	P26664 h genome po	hepatitis			P27960 hepatitis c			h genome	h genome	P26660 h genome po	P26661 h genome po	h genome	ď	0	£	080361 nicotiana t	P21573 xenopus lae	Q02354 saccharomyc	P25502 saccharomyc		Q9ugu0 homo sapien					Q8rh70 fusobacteri	O15530 homo sapien			'n	P23093 plasmodium	۰	P04405 glycine max
SUMMARIES		QI .	POLG_HCV1	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVTW	POLG_HCVJT	2A5D_RABIT	2A5D_HUMAN	RK4_TOBAC	YB1_XENLA	YD49_YEAST	PUT3_YEAST	IF2_LACLC	TF20_HUMAN	TF20_MOUSE	DAPB_STAAM	DAPB_STAAU	DAPB_STAAW	SYN_FUSNN	PDPK_HUMAN	ABF1_STRCX	PUS4_SCHPO	CSP_PLABE	CSP_PLABA	GDF6_BOVIN	GLC2_SOYBN
		DB	-	7	-	П	п	٦	ч	-	7		-	~	-	_	-	٦	-	-	-	-	~	-	~4	-1	-	~	-	-	-	-	-	~	-4
		Length	3011	513	520	520	737	737	3010	3010	3011	3033	3033	3010	3010	586	602	282	303	440	979	950	1960	1983	240	240	240	461	556	825	407	339	347		485
æ	Query	Match	95.5	9.68	9.68	9.6	9.68	9.68	9.68	9.68	89.6	89.6	9.68		84.4	35.4	35.4	34.4	34.1	32.5	32.5	31.8	31.8	31.8	31.2		-	-	31.2	31.2	30.8	30.5	30.5	30.5	30.5
		Score	147	138	138	138	138	138	138	138	138	138	138	131	130	54.5	54.5	2	52.5	20	20	49	49	49	4.8				48	4	47.5	47	47	47	47
	Result	No.	٦	7	m	4	Ŋ	9	7	80	6	07	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30		32	33

	P21711 xenopus lae P08855 orvetolagus								
SYN_BRUMA FMR2_HUMAN	MIXI_XENLA ICAL_RABIT	UBCX_PICPA	PII1_MOUSE	PDPK_MOUSE	PDPK_RAT	P532_HUMAN	CAT3_MAIZE	DOT1_YEAST	POLS_EEVV8
		Н	Н	-4	٦	٦	٦	٦	7
548	377	204	332	559	559	1005	496	582	1254
'nά	~ ~	σ.	σ.	6.	o.	σ.	'n	'n.	٥.
30	30.2	53	53	53	53	53	53	53	29
47	46.5	46	46	46	46	46	45.5	45.5	45.5
3.4	36	38	39	40	41	42	43	44	45

## ALIGNMENTS

10-AUG-1992 (Rel. 22, Created) 10-AUG-1992 (Rel. 22, Last annotation update) 10-AUG-1992 (Rel. 22, Last annotation update) 10-AUG-1992 (Rel. 22, Last annotation update) 10-SEP-2003 (Rel. 22, Length update) 11-SEP-2003 (Rel. 22	POLG_HC	POLG_HCV1 ID POLG_HCV1	STANDARD;	RD;	PRT;	3011 AA	AA.		
11-NGG 1992 (Rel. 2) Last sequence update) 15-SEP-2003 (Rel. 42, Last sequence update) 16-SEP-2003 (Rel. 42, L		P26664;	66 (64)	1000					
15-SEP-2003 (Rel. 42, Last annotation update ()  Geome polyproctein (Concains: Cassid protein ()  Envelope glycoprotein ()  Envelope glycoprotein ()  Envelope glycoprotein ()  (GE 3.4.22.); Processe/Heilcase NS3 (P70) (HepadryLin)  (EC 3.4.22.); Processe/Heilcase NS3 (P70) (HepadryLin)  (EC 3.4.22.); Processe/Heilcase NS3 (P70) (HepadryLin)  (EC 3.4.22.); Processe/Heilcase NS3 (P70) (HepadryLin)  NS48 (P27); Nonstructural protein NS5A (P56); Nonstructural protein  NS58 (P50) (P70) (NNA-directed RNA polymerase) (EC 2.7.7.48)].  Viriases: SSRNA positive-strand viruses, no DNA stage; Flaviviridae;  Hepadryliugs  NGBI_TAXID=11104;  (1)  MEDI_HEPADRSC; PubMed-1848704;  MEDI_HEPADRSC; MSY PLAY A ROLE IN THE VIRAL RNA. REPLICATION.  "Genetic Organization and diversity of the hepatich processed of CATIVITY: Hydrolysis of four peptide bonds in the viral processed organization of Markary and Ser or Ala in Pl.  "Genetic Organization and Mediolysis of FubO PROTEINS: PROTEIN CAND MRNA.  "I-RONROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN CAND MRNA.  "I-SOBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID IS A COMPLEX OF PROTEIN CAND MRNA.  "I-SOBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID IS A COMPLEX OF PROTEIN CAND MRNA.  "I-SOBUNT: THE VIRION OF THE ENVELOPE CONSISTS OF TWO PROTEIN CAND MRNA.  "I-SOBUNT: THE VIRION OF THE WILLIAM CAND MRNA.  "I-SOBUNT: THE VIRION OF THE WILLIAM CAND MRNA.  "I-SOBUNT: THE VIRION OF THE VIRION OF THE WILLIAM CAND MRNA.  "I-SOBUNT: THE VIRION OF THE WILLIAM CAND MRNA.  "I-SOBUNT: THE VIRION OF THE WILLIAM CAND MRNA.  "I-SOBUNT: THE MAD GLYCHARD BY A LIFE AND MRNA		01-AUG-1992	(Rel. 23,	Last seq	neuce	updat	e)		
Genome polyprotecian [Contains: Capsid protein C (Core protein) (1922):  Envelope glycoprotein [Contains: Capsid protein (2)  (EC 34.21.98); Professo-faliciae NS3 (PPS5); Envelope glycoprotein E2 (CFC) (CFC) (NS1); Professo-faliciae NS3 (PPS6); Nonstructural protein NS34 (PP) (Hepacititin) (NS3); Professo-faliciae NS3 (PPS6); Nonstructural protein NS34 (PPS); Nonstructural protein NS34 (PPS6); Nonstructural NS4 (PPS6);		15-SEP-2003	(Rel. 42,	Last ann	otatio	pdn u	ate)		
Envelope glycoprotein E1 (GP83) (GP83); Envelope glycoprotein E2 (GF86) (GP70) (NIJ); Protean P1 (GP83); Envelope glycoprotein E2 (GES ) (GP70) (NIJ); Proteanse/helicase NS3 (PP0) (Hepaclytrin) (EC 3.4.21.98); Monstructural protein NS58 (PP6); PM (Monstructural Protein NS58 (Monstructural Protein); PM (Monstructural Protein Prot		Genome polyp	rotein [C	ontains:	Capsid	prot	ein C (	Core protein)	(P22);
(GPO) (GPVO) (GPVO) (GPVO) (GPO) (GP		Envelope gly	coprotein	E1 (GP32	) (GP3	5), B	nvelope	glycoprotein	E2
(EC 3.4.21.9) Forcease/helinase NS3 (PV0) (Hepacivinin)  (EC 3.4.21.98); Monstructural protein NS5A (P4); Monstructural protein NS5A (P6); Monstructural NS5A (P7); Monstructural NS5A (P		0/45) (245)	( ISN (	Protein :	NON : /	struc	tural p	rotein NS2 (Pa	71)
NS48 (P27); NORINITACIDELIA FIDERIN NS5A (P56); NORINITACIDIA DIOCENTO NS58 (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)]  Viriuses: SaRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepatitis C virus (1solate 1) (HVV);  Viriuses: SaRNA positive-strand viruses, no DNA stage; Flaviviridae; HEPACIVIRS: (Solate 1) (HVV);  NCBL TAXID-11104;  (1)  ENGLETAXID-11104;  (1)  ENGLETAXID-11104;  (2)  ENGLETAXID-11104;  (3)  ENGLETAXID-11104;  (4)  ENGLETAXID-11104;  (5)  ENGLETAXID-11104;  (6)  ENGLETAXID-11104;  ENGLETA		(EC 3.4.22.	); Protes	se/nelica	Se NS3	0/4)	) (Hepa	North (	
NSSB (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48) Indeptitis C virus (isolate 1) (HCV).  Viruses; SRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepatitis C virus (isolate 1) (HCV).  Viruses; SRNA positive-strand viruses, no DNA stage; Flaviviridae; NCBI_TAXID=11104;  [1]  SEGUENCE FROM N.A.  MEDLINE-91172035; PubMed-1848704;  Gallegos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Bardley D.W., Kno G., Houghton M.;  Foch OL., Richman K.H., Han J.H., Berger K., Lee C., Dong C., Gallegos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Bradley D.W., Kno G., Houghton M.;  Foch OL., Richman K.H., Han J.H., Berger K., Lee C., Dong C., Gallegos C., Colt D., Weddina-Selby A., Barr P.J., Weiner A.J., Bradley D.W., Kno G., Houghton M.;  Foch OL., Richman K.H., Han J.H., Berger K., Lee C., Dong C., Gallegos C., Colt D., Weddina-Selby A., Barr P.J., Weiner A. Bossible MEMBRANE-RELATED FONCTION.  Bradley D.W., Kno G., Houghton M.;  Froc. Natl. Acad. Sci. U.S.A. 88:2451-2455[1991].  -I- CONTALTION: THE SMALL PROTEINS NSAA, NS2B, NS2B, NS4B ARE HYDROPHOBIC, SUGGESTING A ROLE IN THE VIRAL RNA REBLICATION.  NS3 AND NS5 MAY PLAY R ROLE IN THE VIRAL RNA REBLICATION.  NS3 AND NS5 MAY PLAY R HOLE IN THE VIRAL RNA REBLICATION.  -I- CATALATIC ACTIVITY: Hydrolysis of Eur peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.  -I- CATALATIC ACTIVITY: HOUGHOUS TO PEPTIDAEE FAMILY S29.  -I- SUBMITT HE PROTEIN E. THE RNUCLEOCAPSID IS A COMPLEX PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX PROTEIN C AND NRA.  -I- SUBMITT THE PROTEIN ENTELOPE THE ENTELOPE THREE BORD BY AND		(EC 3.4.21.3	Nonotruot	ructurar	proces.	1 NO 1	( #4 ) Y	Nonstructural	procein
Hepatitis C vins (Iran directed NAM polymerase) (EC 2.7.7.48)]. Hepatitis C vins (Isolate 1) (HVV). Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepativirus NCBL_TAXID-11104; [1] SEQUENCE FROM N.A. Hepativirus SEQUENCE FROM N.A. HEDILNE-91172836; PubMed-1848704; Choo QL., Richman K.H., Han J.H., Berger K., Lee C., Dong C., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegoos C., Colt D., Medina-Selby C., Colt D., Coltaboration Decursor polyprotein, commonly with Asp of Glu in the Viral PROPEIN C AND MRNA.		( /23) G*CN	NODS LIGHT	utar proc	CE III	40	); NC	istructural pr	rotein
VITURES; SSRNA DOSITIONS LIBORATED (NEW).  WITURES; SSRNA DOSITIONS LIBORATED (NEW).  WITURES; SSRNA DOSITIONS LIBORATED (NEW).  WEDLINE-9117286; PubMed-1848704;  WEDLINE-9117286; PubMed-1848704;  WEDLINE-9117286; PubMed-1848704;  WEDLINE-9117286; PubMed-1848704;  WEDLINE-9117286; PubMed-1848704;  Gallegos C., Coit D., Medina Selby A., Barr P.J., Weiner A.J.,  Bradley D.W., Kuo G., Houghton M.;  Bradley D.W., Kuo G., Houghton M.;  FORCE NAIL ACAG. Sci. U. S.A. 88.2451-2455 (1991).		Honat (Foo)	rivo) (RNA	-ullecter	T KNA	or yme	rase) (	SC 2././.48)].	
MEDITARY, STRAN POSITIVE STRAN FULLES, IN DAY STRAYS, IN THE STRAN POSITIVE STRAND POSITION. STRAND POSITION STRAND POSITION STRAND POSITION STRAND POSITION STRAND POSITION. STRAND POSITION STRAND POSITION STRAND POSITION STRAND POSITION STRAND POSITION.		Witness C	Why non-t+t	Orace 1)	. ( 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2		. 414.0		1 4 4 4
WEBLINE-TAILO4;  (1)  WEBLINE-TAILO4;  WEBLINE-91172826; Pubbmed-1848704;  WEBLINE-91172826; Pubbmed-1848704;  WEBLINE-91172826; Pubbmed-1848704;  WEBLINE-91172826; Pubbmed-1848704;  Good OL., Richana N. H., Han J. H., Berger K., Lee C., Dong C.,  Gallegos C., Colt D., Medina-Sebby A., Barr P.J., Weiner A.J.,  Ganchet Corganization and diversity of the hepatitis C virus.";  Ganchet Corganization and diversity of the hepatitis C virus.";  FORCE. Natl. Acad. Sci. U. S.A. 88:2451-2455(1991);		Honeofuture	TITEON WIT	ים ביו מוור	2717	מ' עמ'	S AND D	aye; riavivii	Tage;
NCBL.TRAXID-IIIU04; SEQUENCE FROM N.A. MEDLINE-91172826, PubMed-1848704; Choo QL., Richman K.H., Han J.H., Berger K., Lee C., Dong C., Gallegos C., Cott D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegos C., Cott D., Medina-Selby A., Barr P.J., Weiner A.J., Gallegos C., Cott D., Medina-Selby A., Barr P.J., Weiner A.J., "Genetic organization and diversity of the hepatitis C virus."; "Genetic organization and diversity of the hepatitis C virus."; "Genetic organization and diversity of the hepatitis C virus."; "Genetic organization and diversity of the hepatitis C virus."; "Genetic organization and diversity of the hepatitis C virus."; "Genetic organization and diversity of the hepatitis C virus."; "Genetic organization and diversity of the hepatitis C virus."; "Genetic organization and diversity of the hepatitis C virus."; "Genetic organization and diversity of the hepatitis C virus."; "Genetic organization and diversity of the hepatitis RNA No. 1-CATALYTIT AGILL." PROPERS IN THE VIRACOLATION IN THE VIRACOLATION IN THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN EVELOPE THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF THE SMISS INSISTUATE OF SHALL SISTILLE OF BOIL OF THE SMISS INSISTUATE OF THE		mepacivitus.	,,,,						
SEQUENCE FROM N.A.  MEDLINE-91172826; PubMed-1848704;  MEDLINE-91172826; PubMed-1848704;  MEDLINE-91172826; PubMed-1848704;  Gallegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,  Bradley D.W., Kuo G., Houghton M.;  Bradley D.W., Kuo G., Houghton M.;  "Genetic organization and diversity of the hepatitis C virus.";  Proc. Natl. Acad. Sci. U.S.A. 88 2243-2455(1991).  -I PUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE  HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.  NS3 AND NS5 MAY PLAY A POLE IN THE VIRAL RNA REPLICATION.  -I CATALVITIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr In P1 and Ser or Ala in P1.  -I CATALVITIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  (RNA)(N).  -I CATALVITIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  RNAD GLYCOPROTEIN E. THE ENVELOPE CONSISTS OF TWO PROTEINS:  PROTEIN A AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.  -I SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.  -I SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.  -I SIMILARITY: THE PROTEASE BELONG STO PEPTIDASE FAMILY S29.  -I SIMILARITY: THE PROTEASE BELONG STO PEPTIDASE FAMILY S29.  -I SIMILARITY: THE PROTEASE BELONG STO PEPTIDASE FAMILY S29.  -I SIMILARITY: THE PROTEASE BELONG STO PEPTIDASE FAMILY S29.  -I SIMILARITY: THE PROTEASE BELONG STO PEPTIDASE FAMILY S29.  -I SIMILARITY: THE PROTEASE BELONG STO PEPTIDASE FAMILY S29.  -I SIMILARITY: THE PROTEASE BELONG STO PEPTIDASE FAMILY S29.  -I SIMILARITY: THE PROTEASE BELONG STO PEPTIDASE FAMILY S29.  -I SIMILARITY: THE PROTEASE BELONG STO PEPTIDASE FAMILY S29.  -I SIMILARITY: THE PROTEASE BELONG STO PEPTIDASE FAMILY S29.  -I SIMILARITY: THE PROTEASE BELONG STO PEPTIDASE PAMILY S20601;		NCBI_TaxID=I	1104;						
SEQUENCE FROM N.A.  MEDILINE—91172226; PubMed=1848704; Choo QL., Richman K.H., Han J.H., Berger K., Lee C., Dong C., Callegos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Bradley D.W., Kuo G., Houghton M.; Genetic organization and diversity of the hepatitis C virus.; "Genetic organization and diversity of the hepatitis C virus.; "Genetic organization and diversity of the hepatitis C virus.; "Genetic organization and diversity of the hepatitis C virus.; "Genetic organization and diversity of the hepatitis C virus.; "Genetic organization and diversity of the hepatitis C virus.; "Genetic organization and diversity organization. Natl. Mall. Proc. Natl. Natl. A ROLE IN THE VIRAD Natl. Proc. Natl.		[1]							
MEDLINE-91172826; PubMed-1848704;  Choo OL., Richman K.H., Han J.H., Berger K., Lee C., Dong C., Callegos C., Colt D., Medina-Selby A., Bair P.J., Weiner A.J., Callegos C., Colt D., Medina-Selby A., Bair P.J., Weiner A.J., Callegos C., Colt D., Medina-Selby A., Bair P.J., Weiner A.J., Saaledy D.W., Kuo G., Houghton M.;  "Genetic organization and diversity of the hepatitis C virus."; Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1991).  "Forcinow THE SMALL PROTEINS NS2A, NS2A, NSAB, AND NSAB ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRARE-RELATED FUNCTION.  "CATALTION THE SMALL PROTEINS NSAB, NSAB, NSAB AND NSAB ARE HYDROPHOBIC, SUGGESTING A ROLE IN THE VIRAL RNA REPLICATION.  "CATALTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral presuresor polyprotesin, commonly with Asp or Glu in the P6 postition, CATALTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA](N).  "I CATALTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA](N).  "I STRAIN NOT THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN C AND MRNA.  "I STRAIN SAISS-FROT entry is copyright. It is produced through a collaboration between the Swiss Institute. There are no restrictions on its work of the Swiss Institution of Bioinformatics and the EMBL outstation the EURDEN BIOInformatics as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/read and email to license@isb-sib.ch).  "ENBL; M6231; AAA45676.1; "  "ENBL; M6231; AAA45676.1; "  "HIGHERDY: DEBD."  "HIGHERDY: DEBD."  "HIGHERDY: DEBD."  "HIGHERDY: DEBD."  "HIGHERDY: DEBD."  "HIGHERDY: DEBD."		SEQUENCE FRO	M.A.						
Callegos C., Colt D., Medina-Selby A., Barger K., Iee C., Dong C., Bradlegos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Bradlegos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Bradlegos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Bradlegos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J., Bradlegos C., Colt D., Medina-Selby D.W., Kuo G., Hudghton M.; Genetic Corganization and diversity of the hepatitis C virus."; Proc. Natl. Acad Sci. U.S.A. 88:241-2455(1991).  1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC. SUGGESTING A POSSIBLE MEMBRARE-RELATED FUNCTION.  1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS5 MAY PLAY A ROLE IN THE VIRA A ROLE IN THE VIRAD PROTEINS.  1- CATALYTIC ACTIVITY: Hydrolysas of four peptide bonds in the P6 postition, Cys or Thr in Pland Ser or Ala in Pl. CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA] (N).  1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA] (N).  1- SUBUNIT: THE VIRTON OF THIS VIRUS IS A NUCLEOCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN MAND GLYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.  1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institutions as long as its content is in no way modified and this statement is not removed "Usage by and for commedities a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  1- SHEL; MGSJ31; AAA45676.1;		MEDLINE-9117	2826; Pub	Med-18487	04;				
Gallegos C., Cott D., Medina-Selby A., Barr P.J., Weiner A.J., Bradley D.W., Kuo G., Houghton M.;  "Genetic organization and diversity of the hepatitis C virus.";  "Genetic organization and diversity of the hepatitis C virus.";  "Froc. Natl. Acad. Sci. U.S.A. 88:2431-2455(1931).		Choo QL.,	Richman K	.H., Han	J.H.,	Berge	r K., L	se C., Dong C.	•
Bradley D.W., Kuo G., Houghton M.;  Genetic organization and diversity of the hepatitis C virus.";  Genetic organization and diversity of the hepatitis C virus.";  Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1991).		Gallegos C.,	Coit D.,	Medina-S	elby A	Ba	rr P.J.	Weiner A.J.	
*Genetic organization and diversity of the hepatitis C virus."; Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455[1991]. Proc. Natl. Prof. Prof. Natl. Asp or Glu in the PG for Catal. Natl. Natl. Natl. Natl. Natl. Asp or Glu in the PG for Catal. Natl. Na		Bradley D.W.	Kuo G.	Houghton	Σ.		· ·		
Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1991).		"Genetic ord	anization	and dive	raite	of th	tenad a	Trie Cattine	
FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC. SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.  NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.  NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.    CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.		Droc Natl	ייטט העטק	2 2 2	98.245	1-245	7 1001	. CT 2 C A TT (13.	
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.  NSJ AND NSS MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.  NSJ AND NSS MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.  I CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in P1.  I CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).  I CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).  I ROPORTEIN ENVELOPE THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.  I SUBLARITY: THE PROTESE BELONGS TO PEPTIDASE FAMILY S29.  THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its would be an email to license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  EMBL; MG321; AAA45676.1; PROMISH SEED.  MEROPS; U39,001; "MCV-98.  MEROPS; UPROMISH SHAD.		-1 - FINCETON	יי שעה כעא	11 000000	NG NGO	77.7	TEAT OC	SOK OFFICE ONK	
NETATION OF THE VIREAL ENGINE OF THE VIREAL STATES TO THE VIREAL STATES THE THE VIREAL STATES THE THE VIREAL STATES THE OF THE VIREAL STATES OF TWO PROTEINS:  PROTEIN A MAIN GLYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN B. THE VIREAL STATES OF TWO PROTEINS:  PROTEIN AND MAIN.		TONOT :	ייייי טיני	THE PROPERTY AND A SECOND ASSESSMENT OF THE PROPERTY OF THE PR	TOUCH	,	ZB, NS#	A AND NS45 AKE	
-I-CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the precursor polyprotein, commonly with Asp or Glu in the precursor polyprotein, commonly with Asp or Glu in the precursor polyprotein, commonly with Asp or Glu in the precursor polyprotein, commonly with Asp or Glu in the precursor polyprotein, commonly with Asp or Glu in the precursor polyprotein or The In Pland Ser or Ala in Pl'		MIDROPHO NG2 NW	101C, 20GG	ESTING A	FOSSIB T T T	35	MBKANE"	KELATED FUNCT	ION.
precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in P1.			A IAM CON	LAI A ROL		HE VI	KAL KNA	REPLICATION.	
precursor potyprotein, commonsy with Asp or Giu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.  -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA](N).  -!- SIBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN MAND GLYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN AND MINA!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation between the Swiss Institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  PIN: A39166; GNAVC3. PDB; 1A1V; 16-FEB-99. MEROPS; US9001; INTERPORS; US9001; INTERPORS; USPO01; INTERPORS; USPONITY INTERPORS INTERPORS INTERPORS INTERPORS INTE			L ACLIVIT	I: HYGEOL	ysts	no i	r pepta	ie bonds in tr	ne virai
POSALION, CYS OF THE IN PI and Ser Or Ala in PI.		precurso	т ротурго	rein, con	Monty	WITH	Asp or	ilu in the Po	
		position	, Cys or	Thr in Pi	and S	er or	Ala in	P1'.	
[RNA] (N)		-!- CATALYTI	C ACTIVIT	Y: N nucl	eoside	trip	hosphat	a= N diphospl	
LIPOPROTEIN THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN C AND MRNA.  1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  EMBL: M62321; Aa445676.1;  PDB: IHBI; 25-NOV-98.  MEROPS: S29.001;  MEROPS: S29.001;  InterPro: IPRO01410; DEAD.									
LIPOPROTEIN ENVELOPE THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPEID IS A COMPLEX OF PROTEIN C AND MRNA.			THE VIRI	ON OF THI	S VIRU	SIS	A NUCLE	CAPSID COVERE	ED BY A
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.  1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY \$29.  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  EMBL, MG2321; Aa45576.1;		LIPOPROT	EIN ENVEL	OPE, THE	ENVELO	PF. CO	SESTSN	TWO PROPETS	. 22
PROTEIN C AND MRNA.  1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  EMBL; MG321; AAA45676.1;  PDB; 1HBI; 25-NOV-98  WERROPS; S29.001;  MEROPS; S29.001;  InterPro; IPRO01410; DEAD.		PROTEIN	M AND GLY	COPPOTETN	HE	N TIN	LEOCA DC	TO TO A COMPLE	. X.
This SHAILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.  This SWISS-PROT entry is copyight. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  ENEL, MGS121, AAA45676.1; PRS 19.  PIR; A39166; GNWVC3.  PDB; 1A1V; 16-FEB-99.  MEROPS; S29, 001; MEROPS; UNIVERSE 19.  MEROPS; US9, 001; MEROPS; UNIVERSE 19.  MEROPS; USP 1001; MEROPS; UNIVERSE 1001; MEROPS; UNIVERSE 1001; MEROPS; US9, 001; MEROPS; US9		PROTEIN	C AND MRN	4		2			;
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  EWBL, M62321; AAA45676.1;			TV. THE D	DOTTO DE	PLONGS	Ę.	PDTTDAC	DAMTIV COD	
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in o way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  EMBL; MG3916; GWWC3. PDB; IHRI; 25-NOV-98 MEROPS; S29.001; - MEROPS; S29.001; - InterPro; IPRO01410; DEAD.						2 :	1001111	remiter of the	
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  EMBL; MG231; AAA45676.1;		This SWISS-P	ROT Pottry	ייייייי אייייי	ight	1+ 10	produc	od through a	ollaboration
the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.sib.ch/announce/or send an email to license@isb-sib.ch).  EMBL; M62321; AAA45676.1; PIR; A39166; GNWVC3. PDB; IALV; 16-FEB-99. PDB; IHRI; 25-NOV-98. MEROPS; S29.001; INTERPRO; IPRO01410; DEAD.		between the	Swigs In	Stitute	FRIO	o form	ation to	and the FMRT	Oitetation
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  EMBL, M62321; Aa445676.1;		the European	Bioinfor	matics In	Stitut	4	here ar	totatat on c	ione on it
modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  EWBL; M62321; AAA45676.1;  EWBL; M62321; AAA45676.1;  PIR; A39166; GNWVC3.  PDB; IARY; I6-FEB-99.  WEROPS; S29.001;  MEROPS; U39.001;  InterPro; IPR001410; DEAD.		non yd esir	-profit	instituti	200	֡֜֜֝֜֜֜֜֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	יייי יייי דרי
entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  EMBL; M62321; AAA44676.1;  PIR; A39166; GNWVC3.  PDB; 1HEI; 25-NOV-98.  MEROPS; S29.001;  MEROPS; U39.001;  InterPro; IPR001410; DEAD.		modified and	this sta	tement is	1	9170	יי ביי ליי	to his and for	DA OH OH AT
or send an email to license@lab-sib.ch).  EMBL, M62321; Aa445676.1; -  EMBL, 39166; GWWC3.  PDB; 1AB1y; 16-FEB-99.  PDB; 1AB1y; 16-FEB-99.  MEROPS; 529.001; -  InterPro; IPR001410; DEAD.		antition and	introe act	topico F	10000	100		Vinni tohooth	ob Appointment
EMBL; M62321; AAA45676.1; - EMBL; M62321; AAA45676.1; - PIR; AA9166; GNWVC3. PDB; 1A1V; 16.FEB-99. PDB; 1HEI; 25-N01; - MEROPS; S29.001; - InterPro; IPR001410; DEAD.		or condens	י לי ניכש	tooppoor	recillen h-ofb	7	י איריף:	AWW. TOD - DID.	cn/announce,
A39166; GNWVC3. 1A1V; 16-FEB-99. 1HEI; 25-NOV-98. PS; K39.001; PPro; IRP001410; DEAD.		בו אבוות מוו ב	מושדד הם ד	TCEIRGETR	D-STD.	Cm)	1		
A3166; GNWC3. 1A1V; 16-FEB-99. 1HEI; 25-NOV-98. PS; S29.001; PS; U39.001; PS; TPRO01410;		LCCCAM TOWN	LJSFRER .						,
1A1V; 16-FEB-99. 1HEI; 25-NOV-98. PS; S29.001; PS; U39.001; PTC: IPRO1410;		TACADA, MOESEL	, AAAA 50,						
JAIV; 16-FEB-99.  JHEI; 25-NOV-98. PS; S29.001; PS; U39.001; rPro; IPRO01410;		PIK; A39166;	GNWVC3.						
1HEI; 25-NOV-98. PS; S29.001; PS; U39.001; TPIO; IPRO01410;		PDB; IAIV; 1	6-FEB-99.						
10;		PDB; 1HEI; 2	5-NOV-98.						
10;		MEROPS; S29.	001;						
10;		MEROPS: 1139	001						
TDD002522		InterPro: 19	- 0	DEAD					
				HOW Cane	7				-

n

```
CELLULAR AMINOPEPTIDASE
CAPSID PROTEIN C (POTENTIAL).
MATIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELLGASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                               POLYPFOCEALM, DEADC, 1.

Core protein, Glycoprotein, Transferase, RNA-directed RNA polymerase, Transmembrane; Coat protein, Envelope protein, Helicase, ATP-binding, Transmembrane; Nonstructural protein, Hydrolase, Serine protease; INII_MET 1 1 REMOVED FROM CARGAL ASSESSED.
                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                          REMOVED FROM CAPSID PROTEIN C BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                            ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                         " I PRO07095; RNA_pol_DS_PS."
" I PRO070994; RNA_pol_DS_PS."
" 101543; HCV_capsid; 1."
" 101542; HCV_capsid; 1."
" 101542; HCV_capsid; 1."
" 101538; HCV_capsid; 1."
" 11538; HCV_NSI; 1."
" 12907; HCV_NSI; 1."
" 10006; HCV_NS3; 1."
" 10006; HCV_NS4; 1."
" 10006; HCV_NS4; 1."
                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                        HCV_NS1.
HCV_NS2.
HCV_NS4a.
HCV_NS4b.
HCV_NS4b.
HCV_NS5b.
HCV_RGRP.
                                                                                                                                                                                       P01001; HCV_NG4b; 1.
P01506; HCV_NS3a; 1.
P00271; helicase_C; 1.
PP00998; Viral RdRP; 1.
PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              327197 MW;
                   IPRO02531;
IPRO02531;
IPRO04109;
IPRO00745;
IPRO01490;
IPRO01868;
                                                                                                                                                                                                                                           DEXDC:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3011 AA;
                                                                                                                                                                               PF01006;
PF01001;
PF01506;
PF00271;
PF00998;
                                                                                                                                                                                                                                          SM00487;
                              InterPro;
InterPro;
InterPro;
                                                                    InterPro;
                                                                              interPro;
                                                                                        InterPro;
                                                                                                             InterPro;
                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                             ProDom;
                                                                                                                    Pfam; P
Pfam; F
                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
                                                                                                                                                 Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                         SMART;
                                                                                                                                                                     Pfam;
                                                                                                                                                                               Pfam;
                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                      CHAIN
CHAIN
CHAIN
CHAIN
```

Gaps 6 Length 3011; Indels ; Db ; 5.4e-13; 0; 95.5%; Sco. ... 92.9%; Pred. No. ... ... 2; Mismatches Score 147; DB Pred. No. 5.4e-Conservative Local Similarity hes 26; Conserv **Ouery Match** Best Loca Matches

ö

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseéisb-sib.ch). 01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
GPROPOPE GIVENT (GPROPE) (GPR "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct MATRIX PROFEIN (POTENTIAL),
MAJOR ENVELOPE PROTEIN E (POTENTIAL)
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL). (POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL) MEDLINE-92230232; PubMed-1314459; Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; CAPSID PROTEIN C (POTENTIAL). N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC. (GLCNAC. (GLCNAC. (GLCNAC. (GLCNAC. (GLCNAC. (GLCNAC. 513 AA (GLCNAC POTENTIAL. N-LINKED N-LINKED N-LINKED N-LINKED N-LINKED N-LINKED N-LINKED N-LINKED N-LINKED EMBL; D10074; BAA00968.1; -InterPro; IPR002522; HCV\_capsid. InterPro; IPR002521; HCV\_core. 115 191 191 383 369 209 233 234 250 417 421 448 SEQUENCE FROM N.A. NCBI\_TaxID-11111; TRANSMEM CARBOHYD NON\_TER CHAIN CHAIN CHAIN 

PKPOKRNORNTNRRPQDVKFPGGGQIVG 28

3

```
520
56499 MW; AA135246CF20D525 CRC64;
  520 5
520 AA;
                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-31644;
                                                                                                                                                                                   POLG_HCVHK
Q01403;
                                                                                                                                                                                                                                                                                                                       Hepacivirus
  NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                        POLG_HCVHK
                                                                                             ò
                                                                                                                      g
                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELULAR AMINOPEPIDASE.

CAPSID PROTEIN C (POTENTIAL).

MATRIX PROTEIN (POTENTIAL).

MAJOR ENVELOPE PROTEIN (POTENTIAL).

MAJOR ENVELOPE PROTEIN (POTENTIAL).

BY SIMILARIT.

N-LINKED (GLCNAC...) (POTENTIAL).

N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                        MEDLINE-93019030; PubMed-1383400;
Abe K., Inchauspe G., Fujisawa K.;
Genomic characterization and mutation rate of hepatitis C virus
'Genomic characterization and mutation rate of hepatitis C virus
1solated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1952).
I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIXCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                           01-JUL-1993 (Rel. 26, Last Sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein EI (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Frament).
Hepatitis C virus (Isolate HCV-476) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           protein.
REMOVED FROM CAPSID PROTEIN C BY THE
                                                     ;
0
                       Score 138; DB 1; Length 513;
Pred. No. 1.5e-12;
3; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Coat protein; Envelope protein;
943F31E3514CDEF3 CRC64;
                                                                                                                                                                        520 AA
                        Query Match
89.6%; Score 138; DE
Best Local Similarity 85.7%; Pred. No. 1.5e
Matches 24; Conservative 3; Mismatches
                                                                              1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                           5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR00252; HCV_capsid.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_env.
IPfam; PF01543; HCV_core; 1.
Ffam; PF01549; HCV_core; 1.
Ffam; PF01559; HCV_env; 1.
FroDom; PD186062; HCV_NS1; 1.
FroDom; PD186062; HCV_NS1; 1.
FroDom; PD186062; HCV_NS1; 1.
IPR0MPF 1.
INIT_MET 1.
IREMOVE
55704 MW;
                                                                                                                                                                                                  (Rel. 26, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D10688; BAA01530.1; -.
                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                115
191
191
191
196
196
196
196
196
196
491
491
491
513 AA;
                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=31643;
                                                                                                                                                                                                01-JUL-1993
01-JUL-1993
                                                                                                                                                                    POLG_HCVH4
Q01404;
                                                                                                                                                                                                                                                                                                        Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                           POLG_HCVH4
                                                                                                                                             RESULT 3
SO
                                                                                                    , Db
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                                                                 Gaps
                                                                                                                                                                                                                                                                                            01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
(GP68) (GP70) (NS1)] (Fragment).
Hepatitis C virus (isolate HCV-KF) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BY SIMILARITY.

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                 ö
                    Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                               Indels
, DB 1; Lens
1.6e-12;
1;
                  Score 138; DB
Pred. No. 1.6e-
3; Mismatches
                                                                                                                                                                                                                                                             520 AA.
                                                                                                                              1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR, J01925; J01925.
InterPro; IPR002523; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002511; HCV_env.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01539; HCV_core: 1.
                  89.6%;
85.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D10687; BAA01529.1; -.
                  Query Match
Best Local Similarity 85.7
Matches 24; Conservative
                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 115
191
191
369
209
234
418
```

```
RESULT 6
                           FT
FT
FT
FT
FT
FT
FT
SO
                                                                                                                                                                                                                                                                                          å
                                                                                                                                                                                                                                                                                                             Q
                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                     A MEDLINE-9233023; PubMed-1314459;
RA MEDLINE-92330232; PubMed-1314459;
RA Tanaka T., Fukuda K., Okada S.I., Yamamoto K., Lizuka H.,
RT Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
RT Momology to reported isolates: Comparative study of four distinct genotypes.;
RI Virology 188:331-341(1992).
CC THOROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION:
RY STAND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC HIDOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS.
CC SUBDERING:
CC PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration the Ewiss Institute of Bioinformatics and the EWBL outstation-use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial nitities requires a license agreement (see http://www.isb-sib.ch/announce/r send an email to licensee@isb-sib.ch).
                                                                                                                                                                                                             01-NUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Cappid protein Correprotein (Contains: Cappid protein Correprotein (Sortians: Cappid protein Correprotein (Sortians: Cappid protein Correprotein (Sortians: Cappid Correprotein Correprotein Correprotein (Sortians NS) and NS2] (Fragment): Matrix
                                                                                                            Gaps
                                                                                                                                                                                                                                                                                  Heparitis C virus (isolate HC-J5) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                         ..) (POTENTIAL).
                (POTENTIAL)
                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Nonstructural protein.
1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                  Length 520;
                                                                                                       1; Indels
                                                    1D2BD0A6FF27349B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein; Glycoprotein; Coat protein; Envelope protein;
        N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
                                                                         Score 138; DB 1;
Pred. No. 1.6e-12;
3; Mismatches 1;
                                                                                                                                                                                                       737 AA
                                                                                                              1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                56476 MW;
                                                                         89.68;
85.78;
                                                                                        24; Conservative
                                                                                                                                                                                               STANDARD;
                                                                              Local Similarity
                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  116
192
384
734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Transmembrane;
INIT_MET
                     CARBOHYD
NON_TER
SEQUENCE
                                                                                                                                                                                            POLG_HCVJ5
CARBOHYD
CARBOHYD
                                                                  Query Match
Best Local
                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
FF FF S
                                                                                                             à
                                                                                                                                 g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-use by mon-profit institute. There are no restrictions on its modified and this statement is not removed. Usage by and for statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to licensee@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                *** MEDLINE-9223032; PubMed-1314459; A Ckamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., A Tanaka T., Fukuda S., Tsuda F., Mishiro S.; Trull-length sequence of a hepatitis C virus genome having poor squency to reported isolates: comparative study of four distinct g gencypes."; Prology 188:331-341(1992).

***Itology 188:331-341(1992).

***Irology 188:331-341(1992).

***Irology 188:331-341(1992).

***Irology 188:331-341(1902).

***Irology 188:31-341(1902).

***Irology 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 22, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus (isolate HC-J7) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                     (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL).
                                                    (POTENTIAL)
                                                                                                                   POTENTIAL)
                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1; Length 737;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81207 MW; 3AF699D82AD501B1 CRC64;
                                                        N-LINKED (GLCNAC-
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 138; DB 1;
Pred. No. 2.3e-12;
3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      737 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPOKRNORNTNRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NSI.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              85.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     89.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11114;
CARBOHYD
CARBOHYD
                                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLG_HCVJ7
P27961;
                                                                             CARBOHYD
                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVJ7
```

us-09-491-146a-29.rsp

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
66-00me polyprotein (Contains and Capsid protein C (Core protein) (P22);
67-00 (GP68) (GP70) (NRI); Protein El (GP32) (GP35); Envelope glycoprotein E2
67-00 (GP68) (GR70) (NRI); Protein F2 (GP68) (GP70) (Hepacivirin)
6. EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P56); Nonstructural protein NS5A (F66); Nonstructural protein NS5A (F66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                       CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
Pfam; PF01542; HCV_core; 1.

Pfam; PF01539; HCV_env; 1.

Pfam; PF01560; HCV_NS1; 1.

ProDom; PD186062; HCV_NS1; 1.

ProDom; PO186062; HCV_NS1; 1.

ProDom; PO186062; HCV_NS1; 1.

ProDom; PF01560; HCV_NS1; 1.

ProDom; PF01560; HCV_NS1; 1.

ProDom; PF01560; HCV_NS1; 1.

Proportion; PROVED FROM CAPSID PROTEIN C BY THE INIT_MET
                                                                                                                                                                                                                                                                   (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-9623522; PubMed-8647104; Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.; *Non-structural protein 3 of hepatitis C virus inhibits
                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 737;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                      81691 MW; 67DFAE11854122F2 CRC64;
                                                                                           CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                    GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB 1;
Pred. No. 2.3e-12;
3; Mismatches 1
                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                         N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-91140698; PubMed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                               89.68;
85.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24; Conservative.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                2333
3053
3055
4417
4423
4448
534
534
558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SECUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                                                                          649
737
737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
SEQUENCE
                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                    CHAIN
                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 7
POLG_HCVBK
                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Loc
Matches
 ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.lsb-sib.ch/announce/or send an email to license@isp-sib.ch).
                                                                                                                         X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDINE=97015088; PubMed=6861916;
LOVE R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
The crystal structure of hepatitis C virus NS3 proteinase reveals a
trypsin-like fold and a structural zinc binding site.";
                                                                                                                                                                                                                                                                                                          SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROPEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                          precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
                                                                                                                                                                                                                                                                                                                                                         SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR007095; RNA_pol_DS_PS. IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002531, HCV_NS1.
IPR002518, HCV_NS2.
IPR0004109, HCV_NS3.
IPR000745, HCV_NS4a.
IPR001490, HCV_NS5a.
IPR002868, HCV_NS5a.
IPR002868, HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01543; HCV_capsid;
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS3; 1
HCV_NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_env;
HCV_NS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1GX6; 10-APR-02.
1QUV; 26-JUN-00.
8OHM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1JXP; 14-JAN-98.
1NS3; 08-APR-98.
1C2P; 15-NOV-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1A10; 25-MAR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-NOV-99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09-APR-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GNWVTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01006;
                                                                                                                                                                                                                                                                                                 (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PDB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PDB
```

1188 1197 1202 1204 1688

1187 1189 1198 1203 1680

3010 AA;

```
STRAND
         TURN
STRAND
                        HELIX
                                 TURN
                                                                                                                                RESULT
         SHEFF
                                                                                                     q
                                                                  REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPPIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2/E0 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-LIRECTED RNA POLYMERASE (POTENTIAL).
                                    Polyprotein; Giycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                         (POTENTIAL).
                                                                                                                                                                                                            (POTENTIAL).
                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                    POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                               GLCNAC.
                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SWART; SM00487; DEXDC; 1.
                                                                                                                                                                                                         209
234
250
250
305
305
441
442
532
532
532
536
623
                                                                                                                                                                                                                                                                                                        645
2041
2077
2240
2529
                                                                                                                                                                                                                                                                                                                                                                                 074
076
081
085
                                                                                                                                                                                                                                                                                                                                                                                                                                             1103
                                                                                                                                                                                                                                                                                                                                                                                                                097
097
097
                                                                                                                                                                                                                                                                                                                                                   1031
1039
1050
1059
1068
1075
                                                                                                                                                                                                                                                                                                                                                                                                       1085
1086
1090
1093
1093
1109
1109
11120
11120
                                                            3D-structure.
INIT_MET
                                                                                                                                                  TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
STRAND
                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                         CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                          HELIX
STRAND
STRAND
STRAND
TURN
                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                              TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
STRAND
STRAND
STRAND
TURN
STRAND
                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAND
                                                                                                                                                                                                                                                                                                                                                                                                        HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HELIX
TURN
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                             TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                    TURN
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its modified and this statement is not removed. Usage by and for commercial or send an email to licensedisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                     Ol-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
GF68) (GF70) (NS1); Protein F7; Nonstructural protein E2
(GF68) (GF70) (NS1); Protein F7; Nonstructural protein NS2 (P21)
(EC 3-4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
NS4B (P27); Nonstructural protein NSA (P4); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Japanese) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      in the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (RNA)(N).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN A AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
MEDLINE-91088550; PubMed-2175903;
Kato N., Hilitata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohon K.;
"Molecular cloning of the human hepatitis C virus genome from
Japanese pailents with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                      1; Indels
327189 MW; F8422D5ECCFDFD9C CRC64;
                                                     89.6%; Score 138; DB 1;
85.7%; Pred. No. 1.1e-11;
1ve 3; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                     PRT; 3010 AA.
                                                                                                                                                            1 PKPOKRNORNINRRPODVKFPGGGGIVG 28
                                                                                                                                                                                            EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ.
                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11116;
Query Match
                                                                                                                                                                                                                                                                                                                              POLG_HCVJA
ID POLG_HCVJA
AC P26662;
                                                                                                                                                                                                                                                                                                                                                         HE DESCRIPTION OF THE PROPERTY AND AND DESCRIPTION OF THE PROPERTY AND DESCRIPTION
```

```
1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Prince A.M.;
                                                                                                                                                        POLG_HCVH
P27958:
                                                                                                            RESULT 9
POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE,HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET 1 CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 138; DB 1; Length 3010;
Pred. No. 1.1e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC. . .) (POT MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
                                                                                                                                                                                                                                                                                       InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                          HCV_capsid
                                                                                                                                                    IPR002518; HCV_NS2.
IPR0004109; HCV_NS3.
IPR000745; HCV_NS4a.
IPR001490; HCV_NS4b.
IPR002868; HCV_NS5a.
IPR002166; HCV_NS6a.
                                                                        InterPro; IPR002522; HCV_capsion
InterPro; IPR002521; HCV_core.
                                                                                                                  HCV_env.
HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF00271; helicase_C; 1. Pfam; PF00998; viral_RdRP; 1. ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS4a;
HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS5a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART; SM00487; DEXDC;
                                                                                                                                  IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2077 207
2240 224
2788 278
3010 AA;
                                                        InterPro; IPR001410;
                MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF02907;
Pfam; PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                          PF02907;
                                                                                                                                                                                                               InterPro;
InterPro;
                                                                                                                                    InterPro;
                                                                                                                                                        InterPro;
                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
VP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
ODRAN ODRA
```

Gaps

ö

1; Indels

3; Mismatches

24; Conservative

Best Local Similarity

Matches

```
e are no restrictions on its
s its content is in no way
Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             between the Swiss Institute of Bioinformatics and the FWBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no warmodified and this statement is not removed. Usage by and for commercial
                                                                                                                                          15.52P-2003 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15.52P-2003 (Rel. 42, Last annoctation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21) (EC 3.4.29.98); Nonstructural protein NS4 (P4); Nonstructural protein NS4 (P4); Nonstructural protein NS5 (P56); Nonstructural protein NS5 (P56); Nonstructural protein NS5 (P56); Nonstructural protein NS5 (P66) (P70) (Rna-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (laolate H) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          oligonucleotide: the crystal structure provides insights into the mode of unwinding.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -!- FUNCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-99154321; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (RNA)(N)

-I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEDCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEGCAPSID IS A COMPLEX OF PROTEIN C AND MANA.

-I- PTW: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.

-I- SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY U39.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                              PRT; 3011 AA
Structure 6:89-100(1998).
                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11108;
```

```
Gaps
                                              · ·) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                        N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     llarity 85.7%; Pred. No. 1.1e-11; Conservative 3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPOKRNORNTNRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1627 16
1635 16
1640 16
3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 24; Conserv.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1480
                            CARBOHYD
CARBOHYD
STRAND
TURN
HELIX
TURN
STRAND
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
HELIX
HELIX
HELIX
HELIX
TURN
STRAND
HELIX
HELIX
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
POLG_HCVJ6
ID POLG_HCVJ6
                                                                                                                                                                                                                                                                                                                                                                                                                    TURN
STRAND
STRAND
STRAND
STRAND
TURN
STRAND
HELIX
STRAND
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAND
STRAND
STRAND
                              ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@lsb-sib.ch).
                                          DR EMBL: M64463; AAA45534.1;

DR PRIS, A56814; GNWVCH.

DR PDB; 1ARIV; 15-FBE-99.

DR PDB; 1ARIV; 17-UN-98.

DR PDB; 1ARIV; 17-UN-98.

DR MEMOPS; 23-001.

DR MEMOPS; 23-001.

DR MEMOPS; 103-001.

DR INTEPTO: IPRO01251; HCV_caps1d.

DR INTEPTO: IPRO01251; HCV_caps.

DR INTEPTO: IPRO01251; HCV_NS.

DR INTEPTO: IPRO01251; HCV_NS.

DR INTEPTO: IPRO01251; HCV_NS.

DR INTEPTO: IPRO01268; HCV_NS.

DR INTEPTO: IPRO01269; HCV_NS.

DR INTEPTO: IPRO01269; HCV_NS.

DR INTEPTO: IPRO01269; HCV_NS.

DR Pfam; PPO1543; HCV_caps1d; 1.

DR Pfam; PPO1543; HCV_caps1d; 1.

DR Pfam; PPO1001; HCV_NS.; 1.

DR Pfam; PPO1009; HCV_NS.

DR Pfam; PPO1009; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.

CHARGE RELAY SYSTEM (BY SIMILARITY)
ATP (POTENTIAL)
DECH BOX.

N-LINKED (GLCNAC...) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REMOVED FROM CAPSID PROTEIN C BY CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11711
1972
3011
369
11083
11083
11165
1237
209
209
410
423
4413
4413
4413
4413
532
532
555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
```

Š 3033

STANDARD;

```
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local
                                                                                                                                                                     INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NX OCC OFF DE PARKEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
SOUR SETT THE FETT TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
               01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope 91ycoprotein E1 (GP32) (GP35); Envelope 91ycoprotein E2
(GP68) (GPV0) (NS1); Protease/helicase N3 (P70) (Hepacivirin)
(EC 3.4.22.-); Protease/helicase N3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NSA (P56); Nonstructural protein
NSAB (P27); Nonstructural protein NSA (P56); Nonstructural protein
NSAB (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-76) (HCV).
                                                                                                                                                                                                                                                                                                                                                          MEDLINE-92044440; PubMed-1658196; Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H., Wachida A., Miyakawa Y., Mayumi M.; Wachida A., Miyakawa Y., Mayumi M.; Nucleotide sequence of the genomic RNA of hepatitis C virus isolated irom a human carrier: comparison with reported isolates for conserved
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  and divergent regions.";
J. Gen. Virol. 72:2697-2704(1991).
I. Gen. Virol. 73:2697-2704(1991).
I. FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
I. CATALYTIC ACTIVITY: Hydrollysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
I. CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           [RNA](N)
-1- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGNS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPRO02166, HCYERP
IPRO01650; HCYERP
IPRO01650; HG1Case_C.
IPRO07095; RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D00944; BAA00792.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS4b; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS1; 1
HCV_NS2; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SSP; J01303; J01303.
                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EROPS; S29.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF01006;
PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF02907
                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro
```

```
CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00288; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                              Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET
                                                                                                                                                                                                                                                                 SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                 CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 138; DB 1;
Pred. No. 1.1e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                          ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PKPORKTKRNTNRRPODVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (
N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPOKRNQRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LINKED
                                                                                                                                                                                                                                                                                                                     DECH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-AUG-1992 (Rel. 23, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2038
                                                                                                                                                                                    1010
1619
1866
2017
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2038 203
2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus.
NCBI_TaxID=11115;
                                                                                                                                                                                                                                                    347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVJ8
P26661;
                                                                                                                                                                                                                                                                              ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                      SITE
CARBOHYD
                                                                                                                                                                                                                                         CHAIN
TRANSMEM
```

```
dsj.cz.
```

```
7. 01-APR-1993 (Rel. 25, Created)
7. 10-APR-1993 (Rel. 25, Last sequence update)
7. 15-SEP-2003 (Rel. 42, Last annotation update)
7. 15-SEP-2003 (Rel. 42, Last annotation update)
8. 15-SEP-2003 (Rel. 42, Last annotation update)
9. 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                     SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                            (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89.6%; Score 138; DB 1; Length 3033;
85.7%; Pred. No. 1.1e-11;
tive 3; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                       POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MW; 1A173E7E3381FD1A CRC64;
                                                                                                       CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                          N-LINKED (GLCNAC
                                                                                                                                                              ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                 558
578
578
578
527
627
649
1091
2038
2359
2311
2811
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                            ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                             RANSMEM
                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCVTW
                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local S
Matches 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                       a
                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration the European Bioinformatics of Bioinformatics and the EMBL outstation use by non-profit institute. There are no restrictions on its modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                       Jennotypes.,
Virology 188:331-341(1992).
-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRARR-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in p1 and Ser or Ala in P1'.
-!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
SEQUENCE FROM N.A.
MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
Fukuda S., Tsuda F., Mishiro S.;
                                                                                                                                                                                                                                                                                        SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY PLOPEROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN C AND MRNA.
                                                                         "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CABSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
PROTESSE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein; Giycopotein; Transferase; RNA-directed RNA polymerase; Core protein; Gaat protein; Envelope protein; Helicase; APP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET
                                                                                                                                                                                                                                                                                                                                                            SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Interpro; IPR002522; HCV_capsid.
Interpro; IPR00251; HCV_capsid.
Interpro; IPR002519; HCV_NSI.
Interpro; IPR002519; HCV_NSI.
Interpro; IPR004519; HCV_NSI.
Interpro; IPR004109; HCV_NS2.
Interpro; IPR004109; HCV_NS3.
Interpro; IPR004490; HCV_NS4.
Interpro; IPR002668; HCV_NS4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; D10988; BAA01761.1; -. PIR; A40250; GWWJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HSSP; P27958; 1HEI.
MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
CHAIN
CHAIN
CHAIN
```

Gaps

BY A

Gaps

ö

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modifited and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-92295714; PubMed-1318627;
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Tanaka T., Hijkata M., Ishimura Y., Shimotohno K.;
Nakazawa T., Hijkata M., Ishimura Y., Shimotohno K.;
"Molecular cloning of hepatitis C virus genome from a single Japanese
cartier: sequence variation within the same individual and among
infected individuals.";
Infected individuals...;
I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
66-00me polyprotein (Contains Cotaid protein C (Core protein) (P22);
67-68) (GP76) (NSI); Protein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP76) (NSI); Protein P7; Nonstructural protein NS2 (P21)
6C 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
6C 3.4.21.98); Nonstructural protein NSA (P4); Nonstructural protein NSAB (P27); Nonstructural protein NSAB (P56); Nonstructural protein NSAB (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (Isolate HC-JT) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -1- SUBDATT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                     (POTENTIAL).
                                                                                                                                                                                                    (POTENTIAL).
                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                (POTENTIAL)
                                                  POTENTIAL)
                                                                         POTENTIAL)
                                                                                            (POTENTIAL)
                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                     Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC. . .) (POTI
                                                                                                                                                                                                                                                                                                                                     Score 131; DB 1; L
Pred. No. 1.2e-10;
3; Mismatches 1;
                                           N-LINKED (GLCNAC.
                                                                                                                                                                                                                        N-LINKED (GLCNAC N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3010 AA,
                                                                                                                                                                                                                                                                                                                                                                                                                          85.1%;
85.2%;
                                                                                                                                                                                                                                                                                           327047
                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity bo....
-hos 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                 623
645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=31642;
                                                                                                              576
623
645
2041
2077
2240
2529
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVJT
Q00269;
                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                    CARBOHYD
SEQUENCE
      CARBOHYD
                                                CARBOHYD
                                                                     CARBOHYD
                                                                                            CARBOHYD
                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 13
POLG_HCVJT
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAE PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ProDom; PD186062; HCV_NS1; 1.
SWART; SMO0487; DEXDC: 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (N-LINKED (N-LINKED (N-LINKED)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001410; DEAD.
InterPro; IPR001522; HVC-capsid.
InterPro; IPR002521; HCV-core.
InterPro; IPR002511; HCV-core.
InterPro; IPR002519; HCV-RNS.
InterPro; IPR001519; HCV-NS1.
InterPro; IPR001499; HCV-NS3.
InterPro; IPR001499; HCV-NS3.
InterPro; IPR001499; HCV-NS4b.
InterPro; IPR002166; HCV-NS5b.
InterPro; IPR002166; HCV-RGP.
InterPro; IPR002166; HCV-RGP.
InterPro; IPR002166; HCV-RGP.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV-capsid; 1.
Pfam; PF01543; HCV-capsid; 1.
Pfam; PF01560; HCV-RNS; 1.
                                                                                                                                                                                                                                                                  EMBL; M84754; -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam: PF01560; HCV_NS1: 1.
Pfam: PF01538; HCV_NS2: 1.
Pfam: PF02907; HCV_NS2: 1.
Pfam: PF01006; HCV_NS4a: 1.
Pfam: PF01001; HCV_NS4a: 1.
R Pfam: PF01506; HCV_NS4b: 1.
R Pfam: PF00271; helicase_C: 1.
R Pfam: PF00998; Viral_RGRP: 1.
R-NDAM: PF00998; Viral_RGRP: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       369
1083
1107
1165
1165
1237
1319
196
                                                                                                                                                                                                                                                                                                         PDB; 1N64; 25-FEB-03
PDB; 1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                         GNWVTW
                                                                                                                                                                                                                                                                                                                                                   MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1107
1165
1230
1316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           196
209
233
234
250
305
417
                                                                                                                                                                                                                                                                                         PIR; A40244;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3D-structure.
:NIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

Ø

3010 AA; 326573 MW; 94A1C77435D642BB CRC64;

2002

```
SEQUENCE
                                              Query Match
                                                                                                                                                                      2A5D_RABIT
                                                                                                                                                        RESULT 14
                      g
                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                      REMOVED FROM CABSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NAMA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                              InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
                                                                                            HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                    Pfam: PF01543; HCV_capsid; INA_pol_PS
Pfam: PF01543; HCV_capsid; I.
Pfam: PF01543; HCV_capsid; I.
Pfam: PF01550; HCV_core; I.
Pfam: PF01550; HCV_NS1; I.
Pfam: PF01538; HCV_NS2; I.
Pfam: PF01001; HCV_NS3; I.
Pfam: PF01001; HCV_NS4; I.
Pfam: PF01001; HCV_NS4s; I.
Pfam: PF01505; HCV_NS4s; I.
Pfam: PF00271; helicase_C; I.
Pfam: PF00398; Viral_RdRP; I.
ProDom: PD18662; HCV_NS1; I.
SWART: SMO0487; DFVCO.1; I.
                                                                                                                                                                   InterPro; IPR000745; HCV_NS4a.
InterPro; IPR001490; HCV_NS4b.
InterPro; IPR002868; HCV_NS5a.
InterPro; IPR002166; HCV_RGRP.
EMBL; D11168; BAA01943.1; -
PIR; A45573; A45573.
PDB; 1A10; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
MEROPS; S29.001; -
                                                                                         IPR002522; H
IPR002521; H
IPR002519; H
IPR002531; H
                                                             MEROPS; U39.001; -. InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                          SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           369
1083
11165
11237
1237
1319
209
234
250
305
417
423
4430
540
556
                                                                                                                                        InterPro; IPR002518;
InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1165
1230
1316
                                                                                                                                                                                                                                                                                                                                                                                                                                          3D-structure.
INIT_MET
                                                                                        InterPro;
                                                                                                     InterPro;
                                                                                                               InterPro;
                                                                                                                            InterPro;
                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME, COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa CONTANT REGULATORY SUBUNIT (SUBUNIT A), THAT ASSOCIATES WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE HITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE B (THE RZ/B/FR55/B55, R3/B''/PR72/PR130/PR59 AND R5/B'/B56 AND R5/B'/B56 AND CELL SIGNALING MOLECULES.
--- SUBCELLULAR LOCATION: Nuclear (By similarity).
--- SUBCELLULAR LOCATION: Nuclear (By similarity).
--- SIBCELCULAR LOCATION: Highly expressed in brain.
--- SIBCELCULAR LOCATION: Highly expressed in brain.
--- SIBCELLULAR LOCATION: Highly expressed in brain.
                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                              2A5D_RABIT STANDARD; PRT; 586 AA.

Q28653; Q28655;
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2011 (Rel. 40, Last annotation update)
56-Insthreonine protein phosphatase 2A, 56 kDa regulatory subunit,
Befala isoform (PP2A, B subunit, B' delta isoform) (PP2A, B subunit, B' delta isoform) (PP2A, B subunit, PR61 delta isoform) (PP2A, B subunit, PR61 delta isoform) (PP2A, B subunit, PR61 delta isoform) (PP2A, B subunit, B'-gamma).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SH3 BINDING, CLASS I (POTENTIAL), NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Multigene family; Nuclear.protein; Repeat.
DOMAIN 21 36 8 X 2 AA APPROXIMATE TANDEM REPEATS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Oryctolagus cuniculus (Rabbit).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -1. CAUTON: NOMENCLATURE USED IN REF.1 REFERS TO PP2A B SUBUNIT B' GAMMA ISOFORM, WHICH IS CITED AS PP2A B SUBUNIT DELTA-PR61 ISOFORM IN LATER PUBLICATIONS.
                Length 3010;
Score 130; DB 1; Length 30
Pred. No. 1.6e-10;
3; Mismatches 2; Indels
                                                                                                              1 PKPOKRNORNINRRPODVKFPGGGGIVG 28
                                                                                                                                                  84.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL, U38193, AAC48532.1; --
EMBL, U38195; AAC48534.1; --
InterPro; IPR002554; B56.
Pfam; PF01603; B56; 1.
                      1 Similarity 82.1
23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=9986;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 507
                      Best Local
Matches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                              THE TENT AND DESCRIPTION OF THE PROPERTY OF TH
```

```
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TISSUE-Fetal brain;
MEDLINE-96355607; PubMed-8703017;
MCCright B., Rivers A.M., Audlin S., Virshup D.M.;
"The B56 family of protein phosphatase 2A (PP2A) regulatory subunits encodes differentiation-induced phosphoproteins that target PP2A to
                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                       550-559; 573-580 AND 584-601 (DEUTA-1).
TISSUE-Bone marrow, and Brain cortex;
MEDLINE-96159032; PubMed-856219;
Tanabe O., Nagase T., Murakaml T., Nozaki H., Usui H., Nishito Y.,
Hayashi H., Kagamiyama H., Takeda M.;
"Molecular cloning of a 74-kDa regulatory subunit (B' or delta) of human protein phosphatase 2A.";
                                                                                                                      014738; 000494; 000696; 015171; 000.
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2003 (Rel. 42, Last annotation update)
Serine/threonine protein phosphatase 2A, 56 kDa regulatory subunit, delta isoform (PPAA, B subunit, B' delta isoform) (PP2A, B subunit, B' delta isoform) (PP2A, B subunit, R5 delta isoform) (PP2A, B subunit, R5 delta isoform) (PP2A, B subunit, BC)
                                                                                                                                                                                                                                                                                                                                                           MEDLINE-97324098; PubMed-9180267; Tanabe O., Gomez G.A., Nishito Y., Usul H., Takeda M.; Mulecular heterogeneity of the cDNA encoding a 74-kDa regulatory subunit (B' or delta) of human protein phosphatase 2A."; FEBS Lett. 408:52-56(1997).
                                                                                                                                                                                                           Homo sapiens (Human).
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                  11;
                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A. (ISOFORM DELTA-2), AND SEQUENCE OF 501-508;
                  Length 586;
                                   Indels
68090 MW; E149A309CDDA7495 CRC64;
                                                                   .9
                                                    1 PKPQKRNQ-RNTNRRPQD-----VKFPGGGQIV 27
                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. (ISOFORMS DELTA-1 AND DELTA-3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. (ISOFORMS DELTA-1 AND DELTA-2)
                                   5.
                                                               1;
                  DB
                                                                                                                    602 AA
               Score 54.5; Di
Pred. No. 2.7;
                                  9; Mismatches
                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. (ISOFORM DELTA-1).
                                                                                                                                                                                                                                                                                                                both nucleus and cytoplasm.";
J. Biol. Chem. 271:22081-22089(1996)
                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EBS Lett. 379:107-111(1996).
                  35.4%;
34.2%;
                                    13; Conservative
                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                    TISSUE-Brain cortex;
586 AA;
                          Local Similarity
                                                                                                                                                                                                                                        NCBI_TaxID=9606;
                                                                                                                   2A5D_HUMAN
SECUENCE
                  Query Match
                                                                                                           2A5D_HUMAN
                                    Matches
တ္တ
                                                                      g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modifited and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; $68686; $68686.

Genew; HGNC:9312; PPP2R5D.

GK; O14738; -.

MIM; 601646; -.

GO; GO:0005634; C:nucleus; TAS.

GO; GO:0007399; P:neurogenesis; TAS.

InterPro; IPR002554; B56.

Pfam; PF01603; B56; I.
                                                                                                                                                                                                                                            å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
MASSING (In 180form Delta-3).
/FTIGH-VSP_005110.
Missing (in 1soform Delta-2).
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
-!- FUNCTION: THE B REGULATORY SUBMYIT MIGHT MODULARE SUBSTRATE SELECTIVITY AND CALALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ISOId=Q14738-3; Sequence=VSP_005110; TISSUE SPECIFICITY: ISOFORM DELTA-2 IS WIDELY EXPRESSED. ISOFORM DELTA-1 IS HIGHLY EXPRESSED IN BRAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -!- SIMILARITY: Belongs to the phosphatase 2A regulatory subunit B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8 X 2 AA APPROXIMATE TANDEM REPEATS
                                                                                                                                                                                                                             PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nuclear protein; Phosphorylation; Alternative splicing; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SH3 BINDING, CLASS I (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1; Length 602;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -1- INDUCTION: BY RETINOIC ACID; IN NEUROBLASTOMA CELL -1- PTM: AT LEAST ISOFORM DELTA-1 IS PHOSPHORYLATED ON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F15F71AF4E565387 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Event-Alternative splicing; Named isoforms-3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /FTId-VSP_005111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 54.5; DE
Pred. No. 2.7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IsoId=Q14738-2; Sequence=VSP_005111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IsoId-Q14738-1; Sequence-Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AB000634; BAA20381.1; -. EMBL; AB000635; BAA20382.1; -. EMBL; D78360; BAA11372.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL, BC001095; AAH01095.1; -. EMBL, BC001175; AAH01175.1; -. EMBL; BC010692; AAH10692.1; -. PIR; S68686; S68686.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                69991 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; L76702; AAB69751.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  35.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              530
565
116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             602 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Name-Delta-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Name-Delta-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Name-Delta-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            523
548
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               85
                                                                                                                                                                                                     COMPARTMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -1- PTM: AT L
RESIDUES.
                                                                                                                                                                                                                                  SUBUNIT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Multigene DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
```

Matches 13; Conservative 9; Mismatches 5; Indels 11; Gaps

1 PKPQKRNQ-RNTNRRPQD------VKFPGGGQIV 27 |:|| ::| ::|:|| : 44 PQPQAQSQPPSSNKRPSNSTPPPTQLSKIKYSGGPQIV 81

Search completed: August 7, 2003, 11:20:02 Job time : 5.90909 secs

Ω

ó

000000000000000000000

hepatitis hepatitis

hepatitis hepatitis

hepatitis

08jyr6 | 068712 | 068712 | 068708 | 068679 | 068683 | 068708 | 068708 | 068708 | 068708 | 098751 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098757 | 098

hepatitis hepatitis hepatitis

hepatitis hepatitis

hepatitis

hepatitis hepatitis hepatitis

089456 099456 089456 089964 089955 089960 089957

hepatitis

hepatitis hepatitis

hepatitis

hepatitis hepatitis hepatitis

hepatitis hepatitis

hepatitis hepatitis hepatitis

```
STRAINFWLESS;
WEDLINE-95062197; Pubmed-7972001;
A Tokita H., Okamoto H., Tsuda F., Song P., Nakata S., Chosa T.,
Tokita H., Mishiro S., Miyakawa Y., Mayuni M.;
A izuka H., Mishiro S., Miyakawa Y., Mayuni M.;
Ilzuka H., Mishiro S., Miyakawa Y., Mayuni M.;
Thepatitis C virus variants from vietnam are classifiable into the seventh, eighth, and ninth major genetic groups.";
Proc. Natl. Acad. Sci. U.S.A. 91:11022-11026(1994).
C. ILPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN CAND MRNA (BY SIMILARITY).
C. PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN CAND MRNA (BY SIMILARITY).
C. EMBL. D88467; BAA13616.1;
C. EMBL. D88467; BAA13616.1;
C. INFOROSS21; HCV_core.
DR InterPro; IPRO02521; HCV_core.
DR InterPro; IPRO02531; HCV_core.
DR Pfam; PF01542; HCV_core;
DR Pfam; PF01543; HCV_CORE;
DR 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MX-1997 (TrEMBLrel. 03, Created)
01-MX-1997 (TrEMBLrel. 03, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core, env. and part of E2/NS1 (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               415 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ALIGNMENTS
08JYRB
08JYRB
08JYRC
08JYRC
068612
068610
068610
068610
068610
068610
06870
06870
06870
06870
06870
06870
06870
06870
06870
06870
06870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     080P70
080P54
090T57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q8QP52
Q8QP88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   090T56
080P90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q8QP55
Q8QP60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                               HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
[2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF01560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-VN235;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-VN235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Okamoto H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P87751
P87751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 1

      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DDT REPRESENTATION OF THE PROPERTY OF THE PROP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          hepatitis or hepat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hepatitis or hepat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P87751 hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       hepatitis
                                                                                                                                                                                                                                                           (without alignments)
281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                          August 7, 2003, 11:05:41; Search time 25.6364 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       086772
096983
0991378
0991378
088778
088778
088778
089778
089778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    068306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q8jyr9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       830525
                          GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    830525 segs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                             protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      09E983
09J3H8
09QIY9
08V7V3
08V7W3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08V7W8
08V7W5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q68306
Q68309
Q8JYR9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08JWL7
068311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_human:*
sp_invertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_organelle:*sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                         US-09-491-146A-29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_fund1:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          103
3010
3010
191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1000.0
1000.0
991.8
991.6
990.9
990.9
990.9
899.6
899.6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          141
140
140
140
1138
1138
1138
                                                                                                                                                                                                                                                                                                                                                         Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 146
142
141
141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB
Maximum DB
                                                                                                                                                             OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Database
                                                                                                                                                                                                                             Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ٠.
ي
```

~

```
TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
NON_TER 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            lesions."
                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                     Q8V7Y2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           068137
                                                                                                                                                                                                                                                                                                                                             RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          068137
ID 0
                  NA WAX SO WAX SO WAX
                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuda F., Miyakawa Y.,
Mayumi M.;
                                                                                                                                                       Gaps
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein; Polyprotein; Transmembrane.
NON TER 415 415 415 SEQUENCE 415 AA; 45301 MW; 39B2040A86517D18 CRC64.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                     ô
                                                                                                             Length 415;
                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Okamoto H.;
Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
                                                           45301 MW; 39B2040A86517D1B CRC64;
                                                                                                                                                                                                                                                                                                                                                              01-NOV-1998 (TrEMBLrel. 08, Created)
01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
Hepatitis C virus.
                                                                                                       ; Score 154; DB 12;
; Pred. No. 1.7e-14;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                PRT; 3013 AA.
                                                                                                                                                                               1 PKPOKRNORNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                    PKPQKRNQRNINRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DR InterPro; IPR002518; HCV_NS1.

LICEPPO; IPR004109; HCV_NS3.

DR InterPro; IPR004109; HCV_NS3.

InterPro; IPR001490; HCV_NS4b.

DR InterPro; IPR001490; HCV_NS4b.

InterPro; IPR002166; HCV_NS5a.

InterPro; IPR007065; RNA_DOl_DS_PS.

Pfam; PF01543; HCV_Copsid; 1.

Pfam; PF01543; HCV_Core; 1.

Pfam; PF01539; HCV_Core; 1.

Pfam; PF01539; HCV_Core; 1.

Pfam; PF01539; HCV_Core; 1.

Pfam; PF01539; HCV_NS1; 1.

Pfam; PF01538; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-98378034; PubMed-9714232;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_capsid.
HCV_core.
HCV_env.
                                                                                                   Query Match
Best Local Similarity 100.0%;
Matches 28; Conservative 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00998; Viral_RdRP;
ProDom; PD186062; HCV_NS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; D84263; BAA32665.1;
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-VN235;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-VN235
                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                          092530
                                                                                                                                                                                                                                                                                  RESULT 2
092530
W KW
SO FT
                                                                                                                                                                                 ò
                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                          44446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DR R DD R B DD R DD R
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
SWART; SM00487; DEXDC; 1.
PROSITE, PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
COAT protein; Envelope protein; 01ycoprotein; Nonstructural protein;
Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3013 AA; 328196 MW; C9EE9C0231E86EAF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ⋖
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                    Hepptitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                ő
                                                                                                             100.0%; Score 154; DB 12; Length 3013; 100.0%; Pred. No. 1.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Kato N.; "Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 191;
                                                                                                                                              0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 191 191
191 AA; 20703 MW; 3215979FE3F45EB7 CRC64;
                                                                                                                                                                                                                                                                                          08V7Y2;
01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 94.8%; Score 146; DB 12; 92.9%; Pred. No. 1.2e-13;
                                                                                                                                                                                                                                                                                                                                                       Core protein (Genome polyprotein) (Fragment).
                                                                                                                                           0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Mismatches
                                                                                                                                                                     1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                        5 PKPOKRNORNINRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genome polyprotein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ilarity 92.9%;
Conservative
                                                                                                                      Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. STRAIN-SA6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11103;
```

103 AA; 11641 MW; B59832ACDA5085EA CRC64;

```
nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     interPro;
 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Trerpro
                                                                                                                                                                                                                                                                 0973HB;
                                                                                                                                                                                                                                            0933н8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                           RESULT 6
Q9J3H8
   S
                                                                                                                                                g
                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                            MEDLINE-92279243; PubMed-1317578;
Bukh J., Purcell R.H., Miller R.H.;
'Sequence analysis of the 7, noncoding region of hepatitis C virus.";
Proc. Natl. Acad. Sci. U.S.A. 89:4442-4946(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               K
                                                                                                                                                                                                                                                                         MEDLINE-93376778; PubMed-8396266;
Bukh J., Purcell R.H., Miller R.H.;
*At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative El general proc. of the putative El general proc. Natl. Acad. Sci. U.S.A. 90.8234-8238(1993).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SUBJECT OF THE NUCLEOCAPSID IS A COMPLEX OF THE NUCLEOCAPS OF TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   District And MEDLINE-20440673; PubMed=10970401;
Tokita H., Kaufmann G.R., Matsubayashi M., Okuda I., Tanaka T.,
Tokita H., Kaufmann G.R., Matsubayashi M., Okuda I., Tanaka T.,
Harada H., Mukaide M., Suzuki K., Cooper D.A.;
"Hepatitis C virus Core Mutations Reduce the Sensitivity of a
Fluorescence Enzyme Immunoassay.";
J. Clin. Microbiol. 38:3450-3452(2000).
I. SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY P.
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL. AB039699; BAB12411.1;
EMBL. AB039699; BAB12411.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 191;
         Bukh J., Purcell R.H., Miller R.H.; "Sequence analysis of the core gene of 14 hepatitis C virus genotypes.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   191
20682 MW; OEB8000CF1F902EB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-WAR-2001 (TrEMBLrel. 16, Created)
01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                        Proc. Natl. Acad. Sci. U.S.A. 91:8239-8243(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 142; DB 12;
Pred. No. 4.5e-13;
2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 103 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRKTQRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
MEDLINE-94336721; PubMed-8058787;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Genome polyprotein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 89.39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
NON_TER 191 1
SEQUENCE 191 AA;
                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                            SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein.
NON_TER
NON_TER 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-HC-N3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus.
                                                                                                                                                                                                                                                               STRAIN-SA6;
                                                                                                                                STRAIN-SA6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9E983;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09E983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 5
Q9E983
STERRERES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sato C.;
                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL: AF207755; AAF65945.1; -- EMBL. GenBank/DDBJ databases.

1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MINA, (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C. "Characteristics of hepatitis C viral genome associated with disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROSITE; PS00190; CYTCCHROME_C; 1.
PROSITE; PS5057; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
POlyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3010 AA; 326421 MW; B10996487CD20668 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                     ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                    Score 141, DB 12; Length 103;
Pred. No. 3.3e-13;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                         01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        91.6%; Score 141; DB 12;
89.3%; Pred. No. 1.2e-11;
tive 2; Mismatches 1;
                                                                                                                                                                                                                                                                      PRT; 3010 AA
                                                                                                            HCV_RdRP.
RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4SSP; P26663; 1JXP.
InterPro; IPR000345; CytC_heme_bind.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_caps1d.
HCV_core.
                       91.6%;
ilarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF01543; HCV_capsid;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 91.6
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR007094; RNA_
                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SM00487; DEXDC;
Query Match
Best Local Similarity
Local 25; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     nterPro; IPR007095
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         progression."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-MD14;
```

```
PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                       lesions."
 ~1
                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                      Q8V7W0
Q8V7W0;
                                               08773
                                                                                                                                                                                                                                                                                                                                                                                                                                  Kato N.
                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                          RESULT 9
                                   RESULT
                                                                                                                                                                                                                                                                                                                                  0877W0
             g
                                                 ద
                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                 PROSITE; PS50507; RDRP_POSITIVE; 1, PROSITE; PS50521; RDRP_VIRAL; 1. Coat protein; Envelope protein; Glycoprotein; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                             Nagayama K., Kurosaki M., Enomoto N., Maekawa S.y., Miyasaka Y.,
Tazawa J.i., Izumi N., Marumo F., Sato C.;
"Time-related changes in full-length hepatitis C virus and hepatitis
                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91.6%; Score 141; DB 12; Length 3010;
89.3%; Pred. No. 1.2e-11;
iive 2; Mismatches 1; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        D816D3BBBF14EE46 CRC64;
                                                            Created)
Last sequence update)
Last annotation update)
                                              PRT; 3010 AA
1 PKPOKRNORNTNRRPQDVKFPGGGQIVG 28
       MEDLINE-20013325; PubMed-10544098;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3010 AA; 326366 MW;
                                                          01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-MAR-2003 (TrEMBLrel. 23, Genome polyprotein. Hepatitis C virus.
                                                                                                                                                                         Virology 263:244-253(1999).
                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                      SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 25; Conserv
                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                           NCBI_TaxID=11103;
                                                                                                                                                                                              STRAIN-MD1-2;
                                                                                                                                 STRAIN-MD1-2;
                                                                                                    Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                              0901Y9
                                                                                                                                                                                                                                                                                                                                                                                                                                      SMART;
                                 RESULT 7
ò
                                               TREACCOSE PTT PTD
```

Gaps

ö

Conservative

```
Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

Interpro: IPR00252; HCV_capsid.

Interpro: IPR002521; HCV_core.

Pfam: PF01543; HCV_capsid.

Pfam: PF01542; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     K
                                                                                                                         0807V3;
01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-007-2002 (TrEMBLrel. 22, Last annotation update)
01-007-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses: ssrNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
-!- SUBMITT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; AB062019; BABB3358 1; -.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                  Kato N.;
"Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               90.9%; Score 140; DB 12; Length 191; 89.3%; Pred. No. 8.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       191 191 191 191 191 191 AM; DDEF95E878228E07 CRC64;
                                                                                                               Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Mismatches
                                                                                                               191
1 PKPOKRNORNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPQKKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
les 25; Conserv
                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                   NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
NON_TER 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE-Liver;
```

```
QBJWL7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        068311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8JWL7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   068311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q8JWL7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                        ö
                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AB062011; BAB83350.1; -.

InterPro; PR002522; HCV_capsid.

Pfam; PF01543; HCV_capsid, 1.

Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                           01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-CT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Hepatitis C virus quasispecies in cancerous and non-cancerous lesions.";
                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                 Score 140; DB 12; Length 191;
Pred. No. 8.9e-13;
1; Mismatches 2; Indels C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 140; DB 12; Length 191;
Pred, No. 8.9e-13;
1; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein.
NON_TER 191 191
SEQUENCE 191 AA; 20667 MW; IEEF95EA0E115BFA CRC64;
                    Polyprotein.
NON_TER 191 191
SEQUENCE 191 AA; 20690 MW; DDEF9785E8228E07 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Core protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                    191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5 PKPQKKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                90.9%;
                                                                                                          90.9%;
Pfam; PF01542; HCV_core; 1.
                                                                                                                               Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity 89.3 nes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       rissue-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      lesions."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kato N.;
                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Kato N.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q8V7W5;
                                                                                                                                                                                                                                                                                                                                                  Q8V7W8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08V7W5
                                                                                                                                                                                                                                                                                                       RESULT 10
Q8V7W8
ID Q8V7W
AC Q8V7W
DT 01-MA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8V7W5
Q8V7W5
Q8 DAC
Q9 DAC
Q9 DD
Q1 DD
Q1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                   ð
```

```
"Comparative Sequence Analysis of the Core Protein and Its Frameshift Product, the Frotein, of Hepatitis C Virus Subtype 1D Strains Product, the Frotein, of Hepatitis C Virus Subtype 1D Strains obtained from Patients with and without Hepatocellular Carcinoma."; Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein (Fragment).
Hepatitis C virus type 1b.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AB062014; BAB8335.1; -.

InterPro; IPR002521; HCV.capsid.

Pfam; PF01543; HCV.capsid.

Pfam; PF01542; HCV.capsid.

Pfam; PF01542; HCV.capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                  Score 140; DB 12; Length 191;
Pred. No. 8.9e-13;
1; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191 191 191 191 191 AM; F5CDCB21037433C8 CRC64;
                                                                                                                                                                                                                                                               191 191 20690 MW; DDEF9785E8228E07 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             90.3%; Score 139; DB 12; 85.7%; Pred. No. 1.3e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPQKKTNRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AB077728; BAC02463.1; -.
InterPro; IPR002522; HCV_copsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_copsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                90.9%;
llarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (TrEMBLrel. 01, (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A. STRAIN-L27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus.
NCBI_TaxID=31647;
                                                                                                                                                                                                                                         Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             068311;
01-NOV-1996 (
01-NOV-1996 (
01-OCT-2002 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ogata S.;
                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12555
```

PRT;

PRELIMINARY;

```
virus.", Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-- Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-- SUBMITT: THE VIRIEN OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
-- PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                     Songsivilai S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thal isolates of hepatitis C
                                                                        01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; U23747; AAA65054.1;
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                 STRAIN-HCV-BB10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein.
NON_TER
SEQUENCE 46
                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-HCV-BB5; Songalvilal S., Kanistanon D., Kunkitti R.; Identification and characterisation of Thal isolates of hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
-- PROTEIN M AND GIXCOPROFIEN B. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; U23744; AAA65011;
                                                                                                                                                                                                                                    Songsivilal S., Kanistanon D., Kunkitti R., "Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 138; DB 12; Length 45;
Pred. No. 3.7e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ouery Match 89.6%; Score 138; DB 12; Length 45; Best Local Similarity 85.7%; Pred. No. 3.7e-13; Matches 24; Conservative 3; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    45 AA; 5015 MW; CC527167096AAA81 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   45 45
45 AA; 5030 MW; CC4C21ED236AAA81 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             45 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPOKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        89.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        uuery Match
Best Local Similarity 85./",
"nag 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                   STRAIN-HCV-BB37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
NON_TER
SEQUENCE 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
068306
DD 068306;
DT 01-NOV-
DD 01
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SOTERNA
```

RESULT 15 Q68309

```
ö
                                                                            Gaps
                                                                            ö
                                         Score 138; DB 12; Length 46;
Pred. No. 3.8e-13;
3; Mismatches 1; Indels
46 46
46 AA; 5129 MW; ABDC4C21ED236AAA CRC64;
                                                                                           1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                            Search completed: August 7, 2003, 11:19:01 Job time: 25.6364 secs
                                                       ilarity 85.7%;
Conservative
                                          89.68;
                                                      Similarity
                                  Query Match
Best Local Simil
Matches 24; C
```

Sequence Seq

Sequence

Appl Appl Appli

Sequence Sequence Sequence

36, 100, 100, 100, 100, 100, 100,

Sequence Sequence Sequence Sequence

```
Sequence 29, Application US/08921887

Patent No. 6030771

GENERAL INFORMATION:
APPLICANT: KINDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAM
WUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100.0%; Score 154; DB 3; 100.0%; Pred. No. 4.9e-15;
US-08-485-500-12
US-08-485-500-14
US-08-867-611-2
PCT-US92-06955A-7
US-07-910-76-12
US-08-440-519-12
US-08-444-818-133
US-08-444-818-133
US-08-440-542-36
US-08-440-519-10
US-08-440-519-10
US-08-440-519-10
US-08-440-519-10
                                                                                                                                                                                                                                                                                                                                           US-09-046-604-36
US-08-440-549-10
                                                                                                                                                                                                                                                                                                                        US-09-388-874-2
                                                                                                                                                                                                                                                                                                                                                                                       US-08-850-328-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION: 435

CILING DATE:
CLASSIFICATION: 435

ATTORNEY AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 35,714

REFERENCE/DOCKET NUMBER: 03063-0380

TELECOMMUNICATION INFORMATION:
TELEPAX: 404-818-3790

TELEFAX: 404-818-3799

INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
'FNGTH: 28 amino acids
'FNGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               internal
    TOFOLOGY:
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER: IBM PC
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30303-1769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FRAGMENT TYPE: J
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ) ORGANISM:
US-08-921-887-29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-921-887-29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Appl
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                App
                                                                                                                                                    August 7, 2003, 11:07:41; Search time 10.5455 Seconds (without alignments) 112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                          328717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /cgn2_6/ptodata/1/laa/5A_COMB.pep:*
/cgn2_6/ptodata/1/laa/5B_COMB.pep:*
/cgn2_6/ptodata/1/laa/6A_COMB.pep:*
/cgn2_6/ptodata/1/laa/6B_COMB.pep:*
/cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
                       5.1.6
Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-921-887-29
US-07-681-703B-26
US-08-485-500-26
US-08-485-500-26
PCT-US92-07813-1
US-07-681-703B-20
US-08-407-410B-20
US-08-407-410B-20
US-08-407-410B-18
US-08-444-818-107
US-07-681-703B-18
US-08-444-818-103
US-08-444-818-103
US-08-444-818-103
US-08-805-18
US-08-444-818-103
US-08-444-818-103
US-08-444-818-103
US-08-444-818-103
US-08-444-818-103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-407-410B-16
US-08-485-500-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                     tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PCT-US91-02370-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -07-681-703B-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-07-681-703B-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -US91-02370-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-407-410B-12
US-08-407-410B-14
                                                                                                                                                                                                                                                                                              1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                             328717 segs, 42310858 residues
                       GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                               protein search, using sw model
                                                                                                                                                                                                                                                                                                                                        BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    length: 0
length: 2000000000
                                                                                                                                                                                                                                                 US-09-491-146A-29
154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sed
                                                                                                                                                                                                                                                                        Perfect score:
                                                                                                                                                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Minimum DB
Maximum DB
                                                                                                               OM protein
                                                                                                                                                                                                                                                                                                Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Database
                                                                                                                                                         Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
```

ö

Gaps

ö

Length 28; Indels

```
Ouery Match 95.5%; Score 147; DB 2; Length 35
Best Local Similarity 92.9%; Pred. No. 5.8e-14;
Matches 26; Conservative 2; Mismatches 0; Indels
                                                                                                                                                MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungauh P.
APPLICANT: Moeckli, Randolph
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPOKRNORNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
CITY: Palo Alto
ADDRESSEE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: Fabian, Gary R.
RECISTATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US/08/485,500
                                                                                                                                                                                                                                                                               FILING DATE:
CLASSIFICATION 415
PIGNOR APPLICATION DATA:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 26, Application US/08485500 Patent No. 5843639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEPHONE: (415) 323-8302
INFORMATION FOR SEQ ID NO: 26
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 35 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CURRENT APPLICATION DATA APPLICATION NUMBER: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                       ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
CLASSIFICATION: 435
           STREET: 350 --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                       USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CITY: Palo
STATE: CA
COUNTRY: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      94306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-407-410B-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 4
US-08-485-500-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match 95.5%; Score 147; DB 1; Length 35; Best Local Similarity 92.9%; Pred. No. 5.8e-14; Matches 26; Conservative 2; Mismatches 0; Indels
                                                                                                                                                                                                   GENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungsuh P.
APPLICANT: Moeckli, Randolph
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/681,703B
FILING DATE: 05-APR-1991
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 26, Application US/08407410B
Patent No. 5843636
GENERAL INFORMATION:
REPERSORY
APPLICANT: Reyes, Gregory
APPLICANT: Moeckli, Randolph
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 26
                                              1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPOKRNORNTNRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                    1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: Fabian, Gary R. REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
                                                                                                                              RESULT 2
WS-07-681-703B-26
Sequence 26, Application US/07681703B
Patent No. 5443965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEPHONE: (415) 324-0880
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      35 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                     ZIP: 94306
COMPUTER REDABLE FORM:
MEDIUM TYPE: Flore
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein .-07-681-703B-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
US-08-407-410B-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                    ò
                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
```

Gaps

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                  Length 35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 75;
                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER REACABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/07813
FILING DATE: 19920916
                                                                                                                                                                                                                                                                                                                                                                                                                         CHAD377/A26D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ENERAL INFURMATION.
APPLICANT: Reges, Gregory
APPLICANT: Kim, Jungsuh P.
APPLICANT: Moeckli, Randolph
TITLE OF INVENTION: Hepatitis C Virus Epitopes
                                                Score 147; DB 5;
Pred. No. 5.8e-14;
2; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 147; DB 5;
Pred. No. 1.3e-13;
2; Mismatches 0
                                                                                                                                                       1 PKPOKRNORNTNRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE: Deblinger & Associates
STREET: 350 Cambridge Ave., Suite 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPOKRNORNTNRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                     Sequence 1, Application PC/TUS9207813
GENERAL INFORMATION:
APPLICANT: LESNIEWSKI, RICHARD R.
APPLICANT: LEUNG, TAX K.
TILLE OF INVENTION: HEPATITIS C ASSAY
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; Sequence 20, Application US/07681703B
; Patent No. 5443965
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD
CITY: ABBOTT PARK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATTORNEY/AGENT INFORMATION:
NAME: POREMBSKIP, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4767.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       95.5%;
                                                Query Match 95.5%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ouery Match
Best Local Similarity 92.9°
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              708-937-9556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: peptide PCT-US92-07813-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RY: U.S.A.
60064-3500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ILLINOIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-07-681-703B-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE:
          PCT-US91-02370-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                      RESULT 6
PCT-US92-07813-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 7
                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                              Score 147; DB 2; Length 35;
Pred. No. 5.8e-14;
2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Reyes, Gregory
APPLICANT: Reyes, Gregory
APPLICANT: Rim, Jungsuh P.
APPLICANT: Moeckli, Randolph
APPLICANT: Moeckli, Randolph
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Peter J. Dehlinger
STREET: P.O. BOX 60850
CITY: Palo Alto
COUNTY: CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:

WEDTUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
APPLICATION NATA:
PCT/US91/02370
FILING DATE: 19910405
FILING DATE:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA: US 594,854
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTONREY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REGISTRATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
SEQUENCE (415) 323-8302
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPOKKNKRNTNRRPODVKFPGGGOIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076.41
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: 19910405
CLASSIFICATION: 435.5
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEPHONE: (415) 323-8302
INFORMATION FOR SEQ ID NO: 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 35 amino acids TYPE: AMINO ACID
                                                                                                                                                                                                                                                                                            35 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-485-500-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 26, Applicat GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PCT-US91-02370-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                 LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

ö

Gaps

Palo Alto

```
95.5%; Score 147; DB 2; Length 90; 92.9%; Pred. No. 1.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 20, Application US/08485500
Patent No. 5843639
CENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungsuh P.
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,500
APPLICAL...
FILING DATE: 09-OCT-12.C
ATORNEY/AGENT INFORMATION:
NAME: Fablan, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/POCKET WUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 323-830.2
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 90 amino acids
TYPE: amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4600-076.21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
CLASSIFICATION: 435
PRICA APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
APPLICATION DATA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 46
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEPHONE: (415) 323-8302
INFORMATION FOR SEQ ID NO: 20
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              90 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 9
US-08-485-500-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 147; DB 1; Length 90;
Pred. No. 1.5e-13;
2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                            WEDLING TECHNOLIS COMPATER
COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
SOFEMATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/681,703B
CLASSIFICATION: 435
PRIOR APPLICATION BATA:
APPLICATION NUMBER: US 505,611
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,611
PRIOR APPLICATION NUMBER: US 594,854
FILING BATE: US 505,611
REGISTRATION NUMBER: 33,875
REGISTRATION NUMBER: 33,875
REFERENCE/ODCKET NUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Repair of APPLICANT: Repair of APPLICANT: Rim, Jungsuh P. APPLICANT: Rim, Jungsuh P. APPLICANT: Rim, Jungsuh P. APPLICANT: Simonsen, Christian C. TITLE OF INVENTION: Hepatitis C Virus Epitopes NUMBER OF SEQUENCES: 26
ADDRESSEE: Peter J. Dehlinger STREET: 350 Cambridge Ave., Suite 100
STATE: CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPOKRNORNINRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 20, Application US/08407410B
Patent No. 5843636
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APP-1990
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           95.5%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEPHONE: (415) 324-0880 INFORMATION FOR SEQ ID NO: 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
LENGTH: 90 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
Best Local Similarity 92.9
Matches 26; Conservative
                                              ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy C
COMPUTER: IBM PC COMP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  , MOLECULE TYPE: protein US-07-681-703B-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: amino acid
TOPOLOGY: linear
                        USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
US-08-407-410B-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                     COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

```
Score 147; DB 3; Length 105;
Pred. No. 1.8e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 18, Application US/07681703B
Patent No. 5443955
GENERAL INFORMATION:
APPLICANT: Kin, Jungsuh P.
APPLICANT: Kin, Jungsuh P.
TITLE OF INVENTION: Heptitis C Virus Epitopes
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Deblinger & Associates
STREET: 350 Cambridge Ave., Suite 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFWMARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/444,818 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPOKRNORNINRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US/07/681,703B
FILING DATE: 05-APR-1991
                                                                                                                                                                                                                                                                                                                       CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14 HAR-1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                             PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: US 505,611 FILING DATE: 06-APR-1990
                                      E: Chiron Corporation 4560 Horton Street
                                                                                                                   COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-D
                                                                                                                                                                                                                                                                                                                                                                                                                                                REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 01:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      95.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (508)359-3876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : 105 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: Harbin, Alisa A. REGISTRATION NUMBER: 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: linear MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRIOR APPLICATION DATA:
                   CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES:
                                                     STREET: 4500 ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET: 350 Camb
CITY: Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-07-681-703B-18
                                        ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-444-818-107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                           ö
                                                           ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      95.5%; Score 147; DB 5; Length 90; 92.9%; Pred. No. 1.5e-13; ive 2; Mismatches 0; Indels
                   Length 90;
                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02370
FILING DATE: 19910405
CLASSIFICATION TATA:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR 1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 107, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Citien, David Y.
APPLICANT: Rutter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
                                                                                                                                                                                                                                                                                APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungsuh P.
APPLICANT: Moeckli, Randolph
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
                 95.5%; Score 147; DB 2;
ilarity 92.9%; Pred. No. 1.5e-13;
Conservative 2; Mismatches 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2; Mismatches
                                                                                                    1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                         1 PKPOKRNORNINRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076.41
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                         PCT-US91-02370-20; Sequence 20, Application PC/TUS9102370; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMpatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE: Peter J. Dehlinger
STREET: P.O. BOX 60850
CITY: Palo Alto
STATE: CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE: (415) 323-8302
INFORMATION FOR SEQ ID NO: 20:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     90 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                 TITLE OF INVENTION: HER NUMBER OF SEQUENCES: 26 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                        Best Local Similarity
Matches 26; Conservi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PCT-US91-02370-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-444-818-107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                            g
```

Indels

```
Gaps
                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 147; DB 2; Length 119;
Pred. No. 2.1e-13;
2; Mismatches 0; Indels
       Score 147; DB 2; L
Pred. No. 2.1e-13;
2; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: TBM PC compartible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE:
                                                                                                                                                                                                                                                              APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungsuh P.
APPLICANT: Kim, Jungsuh P.
APPLICANT: Moeckli, Randolph
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epit
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. 2.16
2; Mismatches
                                                                           1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                 5 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: 4600-076.21 TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 18, Application PC/TUS9102370 GENERAL INFORMATION:
                                                                                                                                                                                                             Sequence 18, Application US/08485500
Patent No. 5843639
Query Match 95.5%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO: 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     95.5%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 95.5
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTIC
                                                                                                                                                                                                                                                                                                                                                                                                                      STREET: 350 Cambor CITY: Palo Alto STATE: CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        amino acid
                                                                                                                                                                                       US-08-485-500-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PCT-US91-02370-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                             á
                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                          Score 147; DB 1; Length 119;
Pred. No. 2.1e-13;
2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPREATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Kim, Unugsuh P.
APPLICANT: Moeckli, Randolph
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
                                           1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                        5 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
RESISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 18, Application US/08407410B
Patent No. 5843636
APPLICATION NUMBER: US 59
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEPHONE: (415) 323-8302
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 92.50
These 26; Conservative
                                                                                                                                                                                                                                               ; MOLECULE TYPE: protein US-07-681-703B-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         amino acid
OGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CITY: Palo
STATE: CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-407-410B-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
```

APPLICANT: Reyes, Gregory APPLICANT: Kim, Jungsuh P

```
Ouery Match 95.5%; Score 147; DB 5; Length 119; Best Local Similarity 92.9%; Pred. No. 2.1e-13; Matches 26; Conservative 2; Mismatches 0; Indels
APPLICANT: Moeckli, Randolph APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes NUMBER OF SEQUENCES: 26
CORRESPONDENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Peter J. Dehlinger
STREET: P.O. BOX 60850
CITTY: Palo Alto
STATE: CA
COMPUTRY: USA
LIP: 94306
COMPUTRY: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Petentin Release #1.0, Version #1.25
CUMRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02370
FILING DATE: 19910405
CLASSIFICATION ATPA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
PRIOR APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
APPLICATION NUMBER: US 594,854
FILING DATE: NOFORMATION:
APPLICATION NUMBER: US 594,854
FILING DATE: US-CT-1990
ATTORNEY/AGENT INFORMATION:
RECESTRATION NUMBER: 4600-076.41
FELECOMMUNICATION INFORMATION:
TELECHOME: Fabian, GATY R.
TELECHOME: Fabian, GATY R.
TELECHOME: Fabian, GATY R.
SEQUENCE CHARACTERISTICS:
LENGRATH: 119 amino acids
LENGRATH: 119 amino acids
TYPE: AMINO ACID
TOTAL AND ACID
TOTAL AND ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY: Innear;
MOLECULE TYPE: protein
PCT-US91-02370-18
```

\_\_sarch completed: August 7, 2003, 11:23:52
Job time : 11.6364 secs

ò

ö

Gaps ö

```
August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1. (oral 6. ptodata 1. pubpaa / USO7_PUBCOMB. pep: *
2. (oral 6. ptodata 1. pubpaa / USO7_PUBCOMB. pep: *
3. (oral 6. ptodata 1. pubpaa / USO7_NEW_PUB. pep: *
3. (oral 6. ptodata 1. pubpaa / USO6_PUBCOMB. pep: *
3. (oral 6. ptodata 1. pubpaa / USO6_PUBCOMB. pep: *
3. (oral 6. ptodata 1. pubpaa / USO8_NEW_PUB. pep: *
3. (oral 6. ptodata 1. pubpaa / USO8_NEW_PUB. pep: *
3. (oral 6. ptodata 1. pubpaa / USO8_PUBCOMB. pep: *
3. (oral 6. ptodata 1. pubpaa / USO8_PUBCOMB. pep: *
3. (oral 6. ptodata 1. pubpaa / USO8_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO9_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO9_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO9_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. ptopa 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. ptopa 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. ptopa 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. ptopa 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. ptopa 1. pubpaa / USO10_PUBCOMB. pep: *
4. (oral 6. ptodata 1. ptopa 1. pubpaa / USO10_PUBCOMB. pep: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       451899
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             otal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPOKRNORNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               451899 segs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                     OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-491-146A-29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Searched:
                                                                                                                                                                                                                                                                                                        Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

200 March Length DB 1D 147 95.5 43 15 US-10-268-562-2 147 95.5 122 15 US-10-268-857B-1 147 95.5 3011 9 US-10-268-857B-1 138 89.6 74 15 US-10-367-677-1 138 89.6 97 10 US-09-916-318-10 138 89.6 97 10 US-09-756-875-8 138 89.6 103 10 US-09-921-397-77 138 89.6 113 10 US-09-921-397-78 138 89.6 137 10 US-09-851-138-46 138 89.6 137 10 US-09-81-138-60 138 89.6 166 11 US-09-87-1-132 138 89.6 166 11 US-09-87-1-132	Result	1	Query	Query			
95.5 43 15 US-10-268-562-2 95.5 122 15 US-10-098-87B-1 95.5 3011 9 US-09-916-352-1 95.6 44 15 US-10-368-562-1 89.6 74 10 US-09-916-359-2 89.6 97 10 US-09-913-138-10 89.6 103 10 US-09-921-337-77 89.6 103 10 US-09-921-397-78 89.6 113 10 US-09-921-38-6 89.6 137 10 US-09-921-38-6 89.6 137 10 US-09-951-138-6 89.6 166 11 US-09-891-138-6 89.6 166 11 US-09-891-138-6	2	acore	Match	Length	9 :	ID	Description
95.5 122 15 US-10-098-857B-1 95.5 190 15 US-10-068-562-1 95.5 3011 9 US-09-916-35-2 89.6 44 15 US-10-367-677-1 89.6 91 9 US-09-916-308-1 89.6 10 US-09-758-308-1 89.6 103 10 US-09-758-3777 89.6 103 10 US-09-921-37-77 89.6 113 10 US-09-921-38-46 89.6 137 10 US-09-921-38-6 89.6 138 10 US-09-851-138-46 89.6 166 11 US-09-899-046-152 89.6 166 11 US-09-899-046-152	~	147	95.5	43	15	US-10-268-562-2	
95.5 190 15 US-10-268-562-1 89.6 44 15 US-10-367-677-1 89.6 74 10 US-09-951-138-10 89.6 91 9 US-09-851-138-10 89.6 97 10 US-09-851-138-10 89.6 103 10 US-09-951-397-77 89.6 103 10 US-09-921-397-77 89.6 137 10 US-09-921-338-46 89.6 137 10 US-09-851-138-46 89.6 166 11 US-09-851-138-60 89.6 166 11 US-09-8951-138-60	7	147	95.5	122	15	US-10-098-857B-1	
95.5 3011 9 US-09-916-359-2 89.6 44 15 US-10-367-677-1 89.6 91 9 US-09-851-138-10 89.6 97 10 US-09-951-39-77 89.6 103 10 US-09-951-397-77 89.6 113 10 US-09-951-138-14 89.6 137 10 US-09-951-138-16 89.6 138 10 US-09-851-138-6 89.6 168 11 US-09-899-066-152 89.6 166 11 US-09-891-135-6	e	147	95.5	190	15	US-10-268-562-1	
89.6 44 15 US-10-367-677-1 89.6 74 10 US-09-581-138-10 89.6 97 10 US-09-758-875-8 89.6 103 10 US-09-951-337-77 89.6 103 10 US-09-921-397-77 89.6 113 10 US-09-921-397-78 89.6 137 10 US-09-921-397-78 89.6 138 10 US-09-851-138-46 89.6 166 11 US-09-891-138-60	4	147	95.5	3011	σ	US-09-916-359-2	Sequence 2. Appli
89.6 74 10 US-09-851-138-10 89.6 91 9 US-09-788-308-1 89.6 103 10 US-09-951-397-77 89.6 108 10 US-09-951-338-14 89.6 113 10 US-09-951-338-46 89.6 137 10 US-09-851-138-6 89.6 138 10 US-09-851-138-6 89.6 166 11 US-09-8951-138-6	2	138	89.6	44	15	US-10-367-677-1	Sequence 1. Appli
89.6 91 9 US-09-758-308-1 89.6 103 10 US-09-756-875-8 89.6 103 10 US-09-921-397-77 89.6 113 10 US-09-921-138-14 89.6 137 10 US-09-851-138-46 89.6 138 10 US-09-851-138-6 89.6 166 11 US-09-899-046-152 89.6 166 11 US-09-899-046-152	9	138	89.6	74	10	US-09-851-138-10	Sequence 10. Appl
89.6 97 10 US-09-756-875-8 89.6 103 10 US-09-921-397-77 89.6 113 10 US-09-921-339-78 89.6 137 10 US-09-921-339-78 89.6 138 10 US-09-851-138-60 89.6 166 11 US-09-899-046-152 89.6 166 11 US-09-878-281-152	7	138	89.6	91	σ	US-09-758-308-1	Sequence 1. Appli
89.6 103 10 US-09-921-397-77 89.6 108 10 US-09-921-381-14 89.6 137 10 US-09-851-138-46 89.6 138 10 US-09-851-138-60 89.6 161 US-09-8951-138-60 89.6 166 11 US-09-895-046-152 89.6 166 11 US-09-898-20	æ	138	89.6	97	10	US-09-756-875-8	Sequence 8. Appli
89.6 108 10 US-09-851-138-14 89.6 113 10 US-09-951-397-78 89.6 137 10 US-09-851-138-46 89.6 138 10 US-09-851-138-60 89.6 166 11 US-09-899-046-152 89.6 166 11 US-09-878-281-152	6	138	9.68	103	10	US-09-921-397-77	Sequence 77. Appl
89.6 113 10 US-09-921-397-78 89.6 137 10 US-09-851-138-46 89.6 138 10 US-09-851-138-60 89.6 166 11 US-09-899-046-152 89.6 166 11 US-09-878-281-152	10	138	9.68	108	10	US-09-851-138-14	Sequence 14. Appl
89.6 137 10 US-09-851-138-46 89.6 138 10 US-09-851-138-60 89.6 166 11 US-09-899-046-152 89.6 166 11 US-09-878-281-152	11	138	9.68	113	10	US-09-921-397-78	Sequence 78, Appl
89.6 138 10 US-09-851-138-60 89.6 166 11 US-09-899-046-152 89.6 166 11 US-09-878-281-152	12	138	89.6	137	10	US-09-851-138-46	Sequence 46, Appl
89.6 166 11 US-09-899-046-152 89.6 166 11 US-09-878-281-152	13	138	9.68	138	10	US-09-851-138-60	Sequence 60. Appl
89.6 166 11 US-09-878-281-152	14	138	89.6	166	11	US-09-899-046-152	Sequence 152, App
	15	138	89.6	166	11	US-09-878-281-152	Sequence 152, App

,		,	;	;		
۰۳	ω «	169	Ξ:	899-046-	4	_
7	0 0	169	7	-03-833-046-4	44	_
בר מ	6.	169	Π	-09-878-281	42,	_
6 T3	89.	169	11	-09-878-281	44	_
0 13	83	182	10	-09-929-955	7	-
1 13	89.	182	14	-10-104-966	~	
2 13	89.	191	Ξ.	-09-194-949	'n	
3 13	æ	318	10	-09-851-138		-
13	89.	319	10	-09-851-138	12,	٠,
5 13	80	319	10	-09-851-138	18,	-
6 13	89.	319	Ξ	-09-899-046		٠.
7 13	89.	319	Ξ	-09-899-046	52,	٦,
. 28 138	89.6	319	1	US-09-899-046-54	Sequence 54, Appl	ب
13	89.	319	1	-09-899-046	144,	۵
0 13	88	319	7	-09-878-281	50,	ب.
13	. 689	319	Ξ:	-09-878-281	52,	_
13	686	319	Ξ:	-09-878-281	54,	_
5 T		319	Ξ:	-09-878-281	144,	۵,
# T	90.0	808	10	-09-973-025	50,	_
13	88	808	Ξ	-866-303	50,	_
13	989	808	7	8-03-332-808	50,	٦,
13	83	808	7	8-09-995-860	50,	٦,
8 13	88	2894	10	S-09-941-611	23,	_
13	89.	2894	15	S-10-044-995	23,	_
0 13	89.	2982	15	S-10-259-275	4	_
13	83	3011	σ	-09-742-659-	4, 4	
2 13	88	3011	10	S-09-952-572-	e 9,	
3 13	89.	3011	10	09-929-955	٦,	
13	80	3011	10	S-09-747-419-	20	_
5 13	ω	3011	11	S-09-891-894-		<b>-</b> 1
				ALIGNMENTS		
RESULT 1						
-10-268-562 -10-268-562	2	, 011 007	900			
; Publication	No. US20030108563A1	03010856	53A1	2		
; GENERAL INFORMATION:	RMATION:					
; APPLICANT: Otho-Clinical Diagnost	Otho-Cl1	nical Di	lagn	tics,		
, TITLE OF IN	VENTION:	Reagent	ij.	the simultaneous	detection of HCV core ant	antige
; TITLE OF INVENTION: antibodies	VENTION:	antibo	00 Te	m		
. Olibbens Appl	TOWELOW	VED /		0,000		
CHRRENT PIL	TNG DATE	SOC .	36	06,802/0		
; PRIOR APPLICATION NUMBER: US	CATION N	OMBER: [	35.5	0/347,943		
; PRIOR FILING DATE	3 DATE:	2001-11	7			
Œ,	SEQ ID NOS	NOS: 3				
TWARE:	PatentIn version	version	3.1			
7					٠	٠
. Deck						

ē

ö

Gaps

ö

Length 43; Indels

Score 147; DB 15; Pred. No. 2.2e-13; 2; Mismatches 0;

Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative 2

; TYPE: PRT ; ORGANISM: Hepatitis C virus US-10-268-562-2

LENGTH: 43

 Sequence 1, Application US/10098857B
Publication No. US20030032031A1
GENERAL INFORMATION:
APPLICANT: GOEDER, JOHN A.
TITLE OF INVENTION: USE OF SPECIFIC ANTIBODY TITERS TO PREDICT HEPATIC

US-10-098-857B-1

RESULT 2

g 2

```
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
TITLE OF INVENTION: TREATING AN HCV INFECTION
CURRENT APPLICATION NUMBER: US/10/367,677
CURRENT FILING DATE: 2003-02-19
PRIOR FILING DATE: 1999-09-07
PRIOR FILING DATE: EARLIER APPLICATION NUMBER: EARLIER APPLICATION NUMBER: EARLIER APPLICATION NUMBER: EARLIER FILING DATE: 1999-03-05
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PATENTIN VET: 2.1
SOFTWARE: PATENTIN VET: 2.1
                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                    Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PUBLICATION INFORMATION: AUTHORS: Ogata, N. et al. TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 138; DB 15;
Pred. No. 3.8e-12;
                                                                                                                                                                    95.5%; Score 147; DB 9; 92.9%; Pred. No. 2e-11; live 2; Mismatches 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                           1 PKPQKRNORNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                           1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; Sequence 10, Application US/09851138; Publication No. US20020183508A1 GENERAL INFORMATION: APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/10367677 Publication No. US20030118604A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: PRT
ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; RELEVANT RESIDUES: 2 TO 45 US-10-367-677-1
                                                                                                                                                                 Query Match 95.5
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 85.7
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: JOLIVET, MICHEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAGES: 3392-3396
DATE: 1991
            LENGTH: 3011
TYPE: PRT
ORGANISM: Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 6
US-09-851-138-10
                                                                                                       US-09-916-359-2
                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/10268562

Sequence 1, Application US/10268562

Publication No. US20030108563A1

GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
TITLE OF INVENTION:
TITLE OF INVENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
         FAILURE IN PEOPLE INFECTED WITH HEPATITIS C VIRUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 190;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A VESULY 4

KESULY 4

KESULY 4

KESULY 6

Sequence 2, Application US/09916359

Petent No. US20200201347341

GENERAL INFORMATION:

APPLICANT: Veronique Barban

TITLE OF INVENTION: TREATING C HEPATITIS

FILE REFERENCE: PMCF97-03A

CURRENT APPLICATION NUMBER: US/09/916,359

CURRENT FILING DATE: 2001-07-26

PRIOR APPLICATION NUMBER: 09/388,874

PRIOR FILING DATE: 1999-09-02

PRIOR FILING DATE: 1997-03-06

NUMBER OF SEQ ID NOS: 2

NUMBER OF SEQ ID NOS: 2

SOFTWARE: FESTSEQ for Mindows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 147; DB 15; I
Pred. No. 1.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 147; DB 15;
Pred. No. 6.6e-13;
FILE REFERENCE: 1300-1800.01
CURRENT APPLICATION WUMBER: US/10/098,857B
CURRENT APPLICATION WUMBER: US/10/098,857B
CURRENT FLING DATE: 2002-10-15
PRIOR APPLICATION WUMBER: 09/616,823
PRIOR FILING DATE: 2000-07-14
PRIOR FILING DATE: 1999-07-15
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PATENTIN Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQKRNQRNTNRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPOKRNORNINRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 95.5%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 95.5%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                  ; ORGANISM: Hepatitis C virus US-10-098-8578-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                  LENGTH: 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-10-268-562-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-10-268-562-1
                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
                                                                                                                                                                                                                                                                                                                  SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ٩
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 77, Application US/09921397

Patent No. US/0020151484A1

GENERAL INFORMATION:

APPLICAMY: HYBRIGENICS

TITLE OF INVENTION: SID nucleic acids and polypeptides selected from TITLE OF INVENTION: applications thereof

FILE REFERENCE: B4809A - JAZ

CURRENT APPLICATION UNMER: US/09/921,397

PRIOR APPLICATION NUMBER: EP 00402225.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 97;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 138; DB 10;
Pred. No. 8.9e-12;
               Sequence 8, Application US/09756875
Patent No. US20020150990A1
GENERAL INFORMATION:
APTEL CHANT: PIKE, IAN
TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 701-E Columbia Square
STREET: 555 13th Street, N. W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89.6%; Score 138;
85.7%; Pred. No. 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                              CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-AUG-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                 MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1808
TELECOMMUNICATION INFORMATION:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 85.73
Matches 24; Conservative
                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ) MOLECULE TYPE: peptide US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            unknown
                                                                                                                                                                                                            CITY: Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                      u.s.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR FILING DATE:
                                                                                                                                                                                                                                  STATE: D. C.
                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE
                                                                                                                                                                                                                                                                       ZIP: 20004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 9
US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQ ID NO 77
LENGTH: 103
                                                                                                                                                                                                                                                      COUNTRY:
US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 14114.0349U2
CURRENT APPLICATION HOWBER: US/09/758,308
CURRENT EPILING DATE: 2001-01-10
CURRENT FILING DATE: 2001-01-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 74;
                                                                                                                                       COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match
89.6%; Score 138; DB 10;
Best Local Similarity 85.7%; Pred. No. 6.7e-12;
Matches 24; Conservative 3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89.6%; Score 138; DB 9;
85.7%; Pred. No. 8.3e-12;
                                                                                                                                                                                                                                                                                    PRIOR AFELCATION NUMBER: 08/836,075
FILING DATE: «UNKNOWN)
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 OCT 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY FAGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/POCKET NUMBER: 1NNS:004
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPOKRNORNTNRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ) MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-138-10
                                                                              ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/09758308 Patent No. US20020090607A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 74 amino acids
  STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PatentIn version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1999-07-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 89.6%
Best Local Similarity 85.7%
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION NUMBER:
PRIOR FILING DATE: 1999-
NUMBER OF SEQ ID NOS: 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBERS PSOFTWARE: PSEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROFHYLACTIC, THERAPEUTIC AND DIAGNO:
                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CURRENTING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 6.0 / ASCII Lext output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 22 Jun 1995
                                                                                                                                                                                                                                                                                   Length 113;
                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     89.6%; Score 138; DB 10;
85.7%; Pred. No. 1.3e-11;
iive 3; Mismatches 1;
                                                                                                                                                                                                                                                                               Score 138; DB 10
Pred. No. 1e-11;
3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                               1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPOKRNORNINRRPODVKFPGGGQIVG 28
      FILE REFERENCE: B4809A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: EP 00402225.7
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFWHARE: PATENTIN Ver. 2.1
SEQ ID NO 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 46, Application US/09851138
; Publication No. US20020183508A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY: USA
21P: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                               89.6%;
                                                                                                                                                                                                                ) ORGANISM: Hepatitis C virus US-09-921-397-78
                                                                                                                                                                                                                                                                             Query Match 89.69
Best Local Similarity 85.77
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
Best Local Similarity 85.73
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JS-09-851-138-46
                                                                                                                                                                                            TYPE: PRT
                                                                                                                                                                         LENGTH:
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                             STUTVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE OF INVENTION: SID nucleic acids and polypeptides selected from TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and TITLE OF INVENTION: applications thereof
                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
        Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
Score 138; DB 10;
Pred. No. 9.5e-12;
3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match 89.6%; Score 138; DB 10; Best Local Similarity 85.7%; Pred. No. 9.9e-12; Matches 24; Conservative 3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 0ct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                          1 PKPOKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
                                                                                                                                                                                                                             Sequence 14, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: USA
21P: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 78, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                            AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET: P.O. BOX 4433
ch 89.6%;
1 Similarity 85.7%;
24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CITY: HOUSTON
STATE: TEXAS
Query Match
Best Local Similarity
                                                                                                                                                                                      RESULT 10
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-851-138-14
                                         Matches
                                                                                    à
                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
```

셤

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                              Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 166;
                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SYSTEM: PC-DOS/MS-DOS
Patentin Release #1.0, Version #1.25 (EPO)
                                                                                                                                                                                                                                                                          Score 138; DB 11;
Pred. No. 1.6e-11;
3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 11;
Pred. No. 1.6e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                        1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Search completed: August 7, 2003, 12:01:13 Job time: 14.3636 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
FILING DATE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
                                                   PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
FILING DATE:
INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 152, Application US/09878281; Publication No. US20030032005A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO: 152: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      89.6%;
                                                                                                                                                                                                                                                                            89.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : 166 amino acids
amino acid
                                                                                                                                                                                                                                                               Ouery Match
Best Local Similarity 85.75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24; Conservative
                                                                                                                                                                                       TOPOLOGY: linear;
MOLECULE TYPE: protein US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; MOLECULE TYPE: protein US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                     FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                  TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        of hepatitis C virus diagnosis, prophylaxis and therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 138; DB 10; Length 138;
Pred. No. 1.3e-11;
3; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                         STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: 08/836,075
FILING DATE: CUNKNOWN>
APPLICATION NUMBER: E9 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: E9 9570076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION: New sequences TITLE OF INVENTION: genotypes for NUMBER OF SEQUENCES: 270 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 152, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                               Sequence 60, Application US/09851138; Publication No. US20020183508A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       E: Floppy disk
IBM PC compatible
                                                                                                                                                                             APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                    STUYVER, LIEVEN
                                                                                                                                                                                                                                                           AGENTS
                                                                                                                                                                                                                                                                                                                                        STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89.6%;
llarity 85.7%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 24; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-899-046-152
                                                                            RESULT 13
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1-09-851-138-60
```

ö

g

```
August 7, 2003, 11:05:37; Search time 38.5455 Seconds (without alignments) 115.301 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /SIDSI/gogdata/geneseq/geneseqp-embl/AA1980.DAT:*/SIDSI/gogdata/geneseq/geneseqp-embl/AA1981.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1981.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1982.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1983.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1983.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1983.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1986.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1986.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1986.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/geneseqg-embl/AA1988.DAT:*/SIDSI/gogdata/genese
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1989_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1989_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1990_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1991_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1991_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1991_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1994_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1995_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1995_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1995_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1995_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqy-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqy-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqy-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqy-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqy-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqy-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqy-embl/AA1997_DAT://SIDSI/gcgdata/geneseqy/geneseqy-embl/AA1997_D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            154
1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1107863 segs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                     OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gapop 10.0 ', Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-491-146A-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Title:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

		Description	NC mosaic protein	NC mosaic protein	Hepatitis C virus	NC mosaic protein	Blood transmiscibl	HCV-S1 full-length	CN14 fragment of H	HCV capsid peptide	HCV capsid peptide
		Ω	AAY06678	AAY06683	AAR58593	AAY06675	AAR53417	AAE20477	AAW06487	AAR30687	AAR30688
		BB	20		15	20	15	23	16	14	14
		Length	28 2	28	312	28	3010	3010	36	38	38
æ	Query		100.0	96.1	95.5	93.5	93.5	93.5	92.9	92.9	92.9
		Score	154	148	147	144	144	144	143	143	143
	Result	NO.	-	10	m	4	S	9	7	æ	6

	10		ζ.		38	14	AAR30689 .	HCV capsid peptide
	11	mı			8 .	15	AAR54065	-A, non-B hep
	7.5	4 4 7 (c)	, ,		1 4	5 6	AAW3/380 AAW66083	
	14	. <b>4</b>			4	20	AAY26952	Hepatitis C virus
	15	43	6		4	21	AAY94410	titis
	16	43	٠. د		٠.	21	AAY94409	Human hepatitis C
	7 6	4 4 4 4	, c	٠	ט גר ט גר	7 [	AAK29534 . AAR29535	HCV core-envelope
	19	. 4 . 6			22	13	AAR29536	HCV core-envelope
	50	<b>4</b> .3	3		201	13	AAR29537	HCV core-envelope
	22	143	22.2	יי סיי	57	E E	AAR29542 AAR25121	HCV core-envelope Non-A. Non-B Henat
	23	43	ď		61	13	AAR20770	Peptide VIIE base
	24	6.43			19	16	AAR69545	Anti-HCV antibody
	26	4 4 3 6			7 19	18	AAW01865	MCV core protein p
	27	43			99	15	AAR12597	PT-NANB viral stru
	28	43	oi o		44	17	AAR96530	Hepatitis C virus
	67	4 4 20 4	× 0		0 CC		AAR29160	Antigen pHCal01.
	31	4 <b>4</b>			0 6	17	AAR41422	Antigen phobiot. HCV fragment 1 / I
	32	43	Ċ		20	14	AAR41423	HCV fragment 2 / I
	33	43	'n		80	15	AAR51389	Branched peptide H
	34	43	oi.		06	16	AAR66619	Hepatitis C virus
	35	ω.	∾.		55	22	AAB31695	
	ود 7 د	4 K	, c		. r	14 14	AAK40978 AAB66633	HCV Core protein N
	38	. <b>4</b>			76	207	AAY01624	protein encoded by
	36	43			97	202	AAW30597	Hepatitis C virus
	40	43	3	1	03	23	ABB77253	HCV bait polypept1
	41	ω. Θ.	ς,	٦,	90	21	AAB18537	Protein encoded by
	2 .	<b>4.</b> 4		٦,	90 5	17	AAR96532	⋍.
	4 4 5 4	4 4 0 6			2.4	17	ABB//254 AAR98348	HCV Dail polypepti HCV capsid core pr
	45	43		1 ~	12	13	AAR29530	e-envelop
							ALIGNMENTS	
מממ								
AAY	JLT 1							
2			standard	d; Prot	otein	n; 2	B AA.	
×××	AAY066	78:						
×								
F 3	17-JUN-1	-1999	(fi	irst e	entry	÷		
OE	NC mosaic		protein	n amino		acid	fragment F.	
Χi								
KW MO	sald	prote	in; endc	endonucle	enic	เหต	eptide; nucleocapsid sisted ligation; vac	sid; NC; hepatitis; REAL; vaccination.
SO	Hepatiti	ပ	virus	. 51				
×	W0991050K	506-81						
×								
G \$		-MAR-1999.						
PF	-AUG	-1998;		98WO-US1738	S173	. 92		
AX PR	25-AUG	-1997;	U1	97US-092188	9218	. 181		
AX PA	HSSN)	a sn (	EPT	HEALTH	æ	HUMAN	N SERVICES.	
×	Fields	H.	Khud	Khudvakov	11 2-			
×	) }							
Z X	ĭ; 1	999-2	04671/17	./17.				
E.	New mo	saic p	rote	in, c	ompr	181	g a plurality o	f homologous antigenic
rg F	pept1d detect	eptides from different genotypes etecting hepatitis infection in	m d1 pati	ffere tis i	nt g nfec	enot	ypes of a species in an individual	- useful

the sensitivity,

us-09-491-146a-28.rag

 $x_{x}x_{y}$ 

a

```
artificial gene and the resulting mosaic protein improve the sensiting spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences AAX106673-683 represent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1994-298800/37.
N-PSDB; AAQ70543.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       312 AA;
                                                                                                                28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JP06225770-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-JUL-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10-JUL-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                           09-MAY-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-AUG-1994
                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                       AAR58593;
                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOFU )
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SANW )
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY06675
ID AAY0
                                                                                                                                                                                                                                                                                                                 RESULT 3
                                                                                                                                                                                                                                                                                                                                    AAR58593
                                                                                                                                                                                                                                                                                                                                                                                         888888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                 The invention relates to a mosaic protein, comprising a plurality of antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted are useful for detecting a hepatitis infection in an individual. The mosaic protein is also useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The invention relates to a mosaic protein, comprising a plurality of monologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for
                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                    28;
                                                                                                                                                                                                                                                                                                                                                                    Length
                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                  100.0%; Score 154; DB 20; 100.0%; Pred. No. 2.3e-14; iive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    peptides from different genotypes of a species detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                            NC mosaic protein amino acid fragment K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                                                                                                                                                                                                                                                                              comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY06683 standard; Protein; 28
                   Claim 5; Fig 9; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1999-204671/17.
                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
es 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO9910506-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                04-MAR-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Fields HA,
                                                                                                                                                                                                                                                                                                                                  Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY06683;
                                                                                                                                                                                                                                                                                                                                                                ery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim
                                                                                                                                                                                                                                                                                                                                                                                     ast Loc
Matches
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAQ70543 is a fragment of hepatits C virus (HCV) or non-B hepatits virus (NAMBH) core and envelope one structural regions, encompassing base pairs 349-1285. It codes for AARS8593 an antigen to a structural region of the HCV virus, which can be used in the diagnosis of NANBH patients and the detection of HCV carriers.
                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                   Non-A non-B hepatitis virus antigens; NANBH; hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A nucleic acid fragment coding Non-A Non-B Hepatitis virus antigens - for diagnosis of NANBH and detection of HCV \,
                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
96.1%; Score 148; DB 20; 96.4%; Pred. No. 1.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 147; DB 15;
Pred. No. 2.6e-12;
1; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (TOKR-) ZH TOKYOTO RINSHO IGAKU SOGO KENKYUSHO
                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus 349-1285 fragment antigen
                                               0; Mismatches
                                                                                      1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SANWA KAGAKU KENKYUSHO CO. TONEN CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 1; Page 18; 22pp; Japanese.
                                                                                                                                                                                                                                                  AAR58593 standard; Protein; 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY06675 standard; Protein; 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 95.5%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         93JP-0193104,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KOKU-) KOKUSAI SHIYAKU KK
                     Best_Local Similarity 96.4
Matches 27; Conservative
```

```
Len
                                                                                                                                                                                                                                                                                       G1y
    Location/Qualifiers
                                                                                                                                                                                                                                                                                   label Ser, Arg,
                                                                                                                                                                                                                                                                                                                                                                     Pro,
                                                                                                                                                                                                                                                                                                                             Ala
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Asn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    _{11e}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gln
                                                                                                                                                                                                                                             label- Gly, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ala
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                label Gly, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               G1u
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Met
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Met
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /label= Cys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              abel- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              label- Thr,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    label- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     label Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    label- Glu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .abel- Glu,
                                     label- His,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          label- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     label= Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     label- Lys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           label- Leu,
                                                                                                                     label- Leu,
                                                                                                                                                             label- Asp,
                                                                                                                                                                                                                                                                                                                           label= Thr,
                                                                                                                                                                                                                                                                                                                                                                   label- Gln,
                                                                                                                                                                                                                                                                                                                                                                                                           label= Lys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                   label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           label- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            label- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label= Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             label- Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             label- Val,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 label = Ser,
                                                                                                                                                                                                     label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                label= Ala,
                                                                               label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 1056
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference 1895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 2143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 2917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 2989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference i
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference
      Key
Misc-difference
                                                            Misc-difference
                                                                                                Misc-difference
                                                                                                                                        Misc-difference
                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosalc protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosalc protein and the artificial mosalc protein are useful for detecting a hepatitis infection in an individual. The mosalc gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, and antigen specificity of enzyme immunosacitylity, and antigen specificity of enzyme immunosassays. This provides improved detection of hepatitis C virus. Sequences AXY06633-683 represent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polymerase chain reaction; PCR; amplify; primer; non-A, non-B hepatitis; NANBH; virus; blood transmissible; detection; hepatitis virus; RT-PCR; C100 antibody; HCV RNA; NS5 region.
                                                                                                                   Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 invention relates to a mosaic protein, comprising a plurality of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           93.5%; Score 144; DB 20; Length 28; 92.9%; Pred. No. 5.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NC mosaic protein amino acid fragment C.
                                                                                                                                                                                                                                                                                                                                                                                  (USSH ) US DEPT HEALTH & HUMAN SERVICES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR53417 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Blood transmiscible NANBHV protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Non-A, non-B hepatitis virus
                                                                                                                                                                                                                                                                                                  98WO-US17385.
                                                                                                                                                                                                                                                                                                                                          97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hest Local Similarity 92.9
Aatches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                       Fields HA, Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1999-204671/17.
                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28 AA;
                                                                                                                                                                                                                    WO9910506-A1
                                                                                                                                                                                                                                                                                                  21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                          25-AUG-1997;
                                     17-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  17-JAN-1995
                                                                                                                                                                                                                                                            04-MAR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR53417;
AAY06675;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
```

The

RESULT 5 AAR53417

ö g

```
36 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; AAT45055
                                                                                                                                                                                                                                                                                                                                                                                                                                                        JP06327482-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                        31-JAN-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29-NOV-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         core region
                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic.
                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                     AAW06487;
                            cassette
                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                 RESULT 7
AAW06487
####X % X O O O O O O O O O O O O O O O X %
                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                   This sequence is encoded by the genome of a blood transmissible non-A, non-B hepatitis (NANBH) virus. The cDNA sequence was isolated using the primers given in AAQ63500-35. The amplified fragments are used in the detection of hepatitis virus. The target DNA was isolated from serum of chronically infected NANBH patients who were C100 antibody-positive and HCV RNA (NS5 region) positive. Reverse transcription-PCR and PCR were performed on cDNA and the total human NANBH DNA was constructed from 23 clones.
                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nucleic acid construct; expression cassette; non-coding region; NCR; untranslated region; UTR; anti-viral drug; drug resistance; HCV-S1; Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                               Blood-transmissible non-A non-B hepatitis virus DNA - used for
                                                                                                                                                                                                                                                                                             93.5%; Score 144; DB 15; Length 3010; 92.9%; Pred. No. 6.9e-11; 1ndels 0; 1ve 1; Mismatches 1; Indels 0;
                                                                                                                                                                                                                                                                                                                                   1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                            (MOLE-) INST MOLECULAR & CELL BIOLOGY (EHRL/) EHRLICH G.
                                                                                                                                                                                                                                                                                                                                                                                                  AAE20477 standard; Protein; 3010 AA
                                                                                                                                                                         Claim 1; Page 8-20; 22pp; Japanese.
        /label- Tyr, Cys
                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV-S1 full-length polyprotein.
                                                                                                                                                         detection of hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20-JUL-2001; 2001WO-IL00669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24-JUL-2000; 2000US-220248P
                                                               92JP-0051885
                                                                                 92JP-0051885
                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Tan YH, Lim SP, Lim SG,
                                                                                                                                                                                                                                                                                                         Local Similarity 92.9
nes 26; Conservative
Misc-difference 2990
                                                                                                                    WPI; 1994-163130/20.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 2002-280605/32
                                                                                                                                                                                                                                                                             3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
                                                                                                   (KAEN/) KAENNO K.
                                                                                                                              N-PSDB; AAQ63499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAD33038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200208447-A2.
                           JP06105690-A
                                             19-APR-1994.
                                                               10-MAR-1992;
                                                                                 10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  31-JAN-2002.
                                                                                                                                                                                                                                                                              Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                    AAE20477;
                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                         AAE20477
ID AAE
                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                            XXXX
```

```
The invention relates to nucleic acid construct which comprises an expression cassette including a first polynucleotide region including a 5' non-coding region (NCR) sequence of an RNA virus and at least an N-terminal portion of a coding sequence of RNA virus, a second polynucleotide region including a 3' untranslated region (UTR) sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus and a third polynucleotide region encoding a reporter promoter, flanked by first and second polynucleotide a regions; and a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette. Nucleic acid construct of the invention is useful for detecting the presence of an RNA virus in a cell. It is also useful for screening anti-viral drugg and determining drug
Novel nucleic acid construct useful for detecting the presence of RNA virus, comprises an expression cassette and a promoter operably linked to expression cassette for minus strand RNA transcription of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This sequence is encoded by the oligonucleotide, CN14, and represents the peptide fragment CP14. CP14 is a fragment of the core region of hepatitis C virus (HCV). CP14 may be used in the detection of HCV infection and to raise antibodies against it.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            resistance of an RNA virus. The present sequence is Hepatitis \hat{C} virus (HCV) isolate HCV-S1 full-length polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Detection of hepatitis C virus - using oligopeptide fragment of HCV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  93.5%; Score 144; DB 23; Length 3010; 92.9%; Pred. No. 6.9e-11; indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CP14; core region; hepatitis C virus; HCV; detection;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                           Example 1; Page 70-81; 81pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAW06487 standard; Protein; 36 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CN14 fragment of HCV core region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 1; Page 6; 7pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                93JP-0156026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         93JP-0156026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (IMMO ) IMMUNO JAPAN KK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1995-047903/07.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3010 AA;
```

Matches

ò g RESULT 8

UXXXXEEXE

```
The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope groups. These peptides can be used in immunoassays for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific antibodies.
                                                                                                                                                                      Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-B, non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.9%; Score 143; DB 14; Length 38; 92.9%; Pred. No. 1.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (Updated on 25-MAR-2003 to correct PN fleld.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR30689 standard; peptide; 38 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Todd JA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      91US-0714471.
91US-0718052.
                                                                                                                                                                                                                                                                                                                                                                                                                        92WO-US03635
                                                                                (first entry)
                                                                                                                           HCV capsid peptide No. 24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV capsid peptide No. 25.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 92.9
Matches 26; Conservative
                                                          (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1993-018073/02
                                                                                                                                                                                                                                                                        Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             38 AA;
                                                                                                                                                                                                                                                                                                                        W09222571-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-JUN-1991;
20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                        29-APR-1992;
                                                    25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                        23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Jolley ME,
     AAR30688;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR30689;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 10
AAR30689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                       ö
                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope groups. These peptides can be used in immunoassays for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The sequences given in AAR30665-89 represent fragments of the
  Length 36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 38;
                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Indels
                                                  1;
  Score 143; DB 16;
Pred. No. 1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.9%; Score 143; DB 14;
92.9%; Pred. No. 1.1e-12;
ive 1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (Updated on 25-MAR-2003 to correct PN field.)
                                                  1; Mismatches
                                                                                                    1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                              1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                        ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR30688 standard; peptide; 38 AA
                                                                                                                                                                                                                                                                      AAR30687 standard; peptide; 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (BAXT ) BAXTER DIAGNOSTICS INC
  92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       910S-0714471.
910S-0718052.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92WO-US03635
                                                                                                                                                                                                                                                                                                                                                                      (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 92.99
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV capsid peptide No. 23
                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1993-018073/02
Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  virus - for immu
non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          W09222571-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                      25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Jolley ME,
                                                                                                                                                                                                                                                                                                                        AAR30687;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Seguence
```

ö

Gaps

ö

Indels

RESULT 9
AAR30688
ID AAR3

ð g

;88888888**%**8

9

ö

```
The present sequence represents a Hepatitis C virus (HCV) protein sequence from the disclosure of the present specification. The present specification describes a chimeric HCV peptide antigen which comprises at least 2 peptide epitope regions from the HCV polypeptide core region, 2 peptide epitope regions from the NS3 region and at least 2 peptide epitope regions from the NS3 region. The antigen binds specifically with an antibody produced by a human infected by HCV. The peptide can detect a wide range of HCV infections with high sensitivity.
                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New chimeric peptide antigen derived from hepatitis C virus protein - useful for detecting HCV infections
                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus; HCV; chimeric; antigen; detection; core region; epitope; NS3; NS4; infection.
                                                                                                                                                                                                ö
                                                                                 Hepatitis C virus #4 and #6 genes were isolated (AAQ64068-69). Both genes contain the core, ENV, NSI, NS2 and NS3 regions. A core region fragment is given in AAQ64067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                 DNA coding a Non-A, non-B hepatitis virus antigen - useful for
detecting HCV within serum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92.9%; Score 143; DB 19;
llarity 92.9%; Pred. No. 1.2e-12;
Conservative 1; Mismatches 1;
                                                                                                                                                                  Score 143; DB 15;
Pred. No. 1.1e-12;
1; Mismatches 1;
                                                                                                                                                                                                                     1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                   1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Page 24; 30pp; Japanese.
                                                        Disclosure; Page 11; 22pp; Japanese
                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus C-1 protein 1-43,
                                                                                                                                                                                                                                                                                                                   AAW37380 standard; Protein; 43 AA.
                                                                                                                                                                Ouery Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97JP-0027015.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               96JP-0024045
                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1998-022248/03.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (TOFU ) TONEN CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis c virus
                                                                                                                                       38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JP09278794-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09-FEB-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                         11-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-OCT-1997.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                        Seguence
                                                                                                                                                                                                                                                                                                                                             AAW37380;
                                                                                                                                                                                                                                                                                        RESULT 12
                                                                                                                                                                                                                                                                                                      AAW37380
    ò
                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                         The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capaid protein of the virus and themselves contain epitope groups. These peptides can be used in immunoassays for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and antibodies.
                                                                                                                                                                                                                  Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Non-A, non-B hepatitis virus; NANBHV; hepatitis C virus; HCV; Core; ENV; NS1; NS2; NS3; antigen; detection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 143; DB 14; Length 38;
Pred. No. 1.1e-12;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Non-A, non-B hepatitis virus coreI region fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (SANW ) SANWA KAGAKU KENKYUSHO CO.
(TOFU ) TONEN CORP.
(TOKR-) 2H TOKYOTO RINSHO IGAKU SOGO KENKYUSHO.
                                                                                                                                                                                                                                                                                                                                                                                                                          (Updated on 25-MAR-2003 to correct PN fleld.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR54065 standard; Protein; 38
                                                                                                                                   (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92.9%;
92.9%;
                                                                 92WO-US03635.
                                                                                         91US-0714471.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92JP-0088140.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 92.9
nes 26; Conservative
                                                                                                                                                              Leahy DC,
                                                                                                                                                                                         WPI; 1993-018073/02.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1994-205028/25.
N-PSDB; AAQ64067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                     38 AA;
                                                                                           13-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JP06141870-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12-MAR-1992;
                                                                 29-APR-1992;
                                                                                                          20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    14-FEB-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12-MAR-1992;
                                       23-DEC-1992.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24-MAY-1994
                                                                                                                                                            Jolley ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR54065;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4atches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AARS4065
ID AAR
888888888888888888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

ö

us-09-491-146a-28.rag

```
N-terminus of the hepatitis C virus core protein. The peptides are used to generate monoclonal antibodies or antibody fragments specific for hepatitis C virus (HCV) core protein. The monoclonal antibodies are used for early diagnosis of HCV infections, especially by two-antibody
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 useful for detecting and/or quantifying hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polypeptides that bind to anti-hepatitis C virus antibodies, useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human; hepatitis C virus; HCV; immunogen; antigen; antibody; virucide; hepatotropic; anti-inflammatory; virus detection; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Peptides AAY26949-Y26955 represent peptide epitopes derived from the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human hepatitis C virus core protein N-terminus, residues 2-45.
                               Epitope; hepatitis C virus; core protein; monoclonal antibody;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Jolivet-Reynaud C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Jolivet M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; DB 20;
1.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Paranhos BC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 143; DB Pred. No. 1.2e 1; Mismatches
                                                          diagnosis; infection; sandwich immunoassay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPORKTKRNAHRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 3; Page 10; 19pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY94410 standard; peptide; 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Yvon S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                98FR-0003087
                                                                                                                                                                                                                                                                                                                                       98FR-0003087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99WO-IB01933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0196155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11-SEP-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ď.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Monoclonal antibodies
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dalbon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C virus core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sandwich immunoassay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Piga N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1999-530397/45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         API; 2000-411934/35.
                                                                                                                                                                                                                                                                                                                                                                                               (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      44 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  H
                                                                                                                                                                                                                                                                             09-MAR-1998;
                                                                                                                                                                                                                                                                                                                                       09-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20-NOV-1998;
                                                                                                                                                                  FR2775690-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        02-JUN-2000.
                                                                                                                                                                                                                      10-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Jolivet RC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dalbon P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY94410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY94410

ID AAY9

XX AAY9

XX AAY9

XX II-S

XX II-S

XX HUMB

XX HUMB

XX HUMB

XX HOPB

XX HOPB

XX YOCC

XX
     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The invention relates to a peptide, which is recognised by antibodies against amino acids 2-45 at the N-terminus of the core (or nucleocapsid) possibly to the period of hepatitis C virus (HCV), or its variants. The peptide has a tertiary structure consisting of two alpha-helical fragments, almost compendicular to each other in space, connected by a junction peptide. Excluded are all proteins and peptides comprising, or consisting of, the N-terminal part of p21 (starting from amino acid to x 2). Also new are cc (1) monoclonal or polyclonal antibodies produced using the peptide as an immunogen and (2) complex consisting of the peptide specifically cound to some other molecule, particularly peptide or nucleotide (1) can functionalised aromatic compounds. The peptide is used (1) cas immunogen for generating antibodies and (11) for detecting and complex formation). Antibodies are used correspondingly to detect HCV or related antigens. The peptide and antibodies may also be used to treat or prevent HCV infections. The present sequence represents the N-terminus contracts of the core p21 protein of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               nseq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New peptide from the N-terminus of hepatitis C virus p21 protein containing the immunodominant epitope - and related antibodies, u for diagnosis, treatment and prevention of hepatitis C infection
                                                                                                                                                                                                                                                                        antigenic; nucleocapsid; p21 protein; alpha-helical; immunogen; antibody; hepatitis C virus; epitope; N-terminal fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92.9%; Score 143; DB 19; Length 44; 92.9%; Pred. No. 1.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Penin F;
                                                                                                                                                                                                                      Hepatitis C virus p21 protein N-terminal fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus Core protein amino acids 2-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ladaviere L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAW66083 standard; peptide; 44 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Disclosure; Page 16; 37pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Lacoux X,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98WO-FR00442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97FR-0002878
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY26952 standard; peptide;
                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1998-495793/42
                                                                                                                                                                                                                                                                                                                                                               hepatitis c virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         44 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     05-MAR-1997;
                                                                                                                                                               16-NOV-1998
                                                                                                                                                                                                                                                                                                                                                                                                                WO9839360-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              05-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11-SEP-1998,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Dalbon P,
                                                                                                        AAW66083;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY26952;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
RESULT 13
AAW66083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY26952
ID AAX2
XX AAX
AC AAXZ
XX DT 21-I
```

SXS

g ð

ö

Gaps

ö

Length 44; Indels

diagnosing and preventing hepatitis C infections -

Claim 1; Page 42; 50pp; English.

The present peptide, designated S42G, corresponds to residues 2 to 45 of the N-terminus of the human hepatitis C virus (HCV) core protein. It is an immunodominant region containing conformational type epitopes and linear type epitopes. It manifests an immunoreactivity with all the sera of individuals or blood samples infected with HCV and which possess substituted for homologous amino acros protein. An amino acid may be substituted for homologous amino acids and side chains and peptide bonds may be modified. For example, L-amino acids may be replaced by D-amino sequence and its antiqenic derivatives may be used for detection of hepatitis C virus and for raising antibodies against the virus. 

44 AA; Sequence

Gaps ó 92.9%; Score 143; DB 21; Length 44; 92.9%; Pred. No. 1.2e-12; tive 1; Mismatches 1; Indels Query Match
'est Local Similarity 92.99
:tches 26; Conservative

ö

1 PKPQRKTKRNAHRRPQDVKFPGGGGIVG 28

ò g

Search completed: August 7, 2003, 11:14:07 Job time : 38.6364 secs

```
August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                             283308
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                       al number of hits satisfying chosen parameters:
                                                                                                                                                                                      US-09-491-146A-28
154
1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                      283308 seqs, 96168682 residues
                                                                           OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                      Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                                        Searched:
                                                                                                              Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. PIR\_76:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\*

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

Minimum DB seq length: 0 Maximum DB seq length: 200000000

# SUMMARIES

•	Description	qenome polyprotein																			genome polyprotein		٠	•	pro	genome polyprotein	genome polyprotein	genome polyprotein	hypothetical prote	genome polyprotein
SUMMAKIES	QI	A45573	541353	841355	_	541348	-	S41341	S41370	S41369	S41368	4134	S41344	S41350	₹	S41345	S41347	S41343	S41346	S21471	S12707	PC1284	JQ1925	JQ1926	JH0711	S19876 ·	S18031	S18032	PN0677	300883
	DB	_	ď	~	~	~	~	~	~	~					~	•	•	~	•	•		•	•			•	•		7	7
	Length	· m	108	108	108	108	112	112	114	114	114	115	115	115	115	115	115	115	118	369	441	513	220	523	550	782	782	782	787	874
æ	Query	93.5	ς.	ď	95.9	ď	ď	95.9	ď	ς.	ä	ä	ď	ä	ä	ä	'n	92.9	ς.	ä	ä	'n	'n.	ď	ď.	ä	ä	ς.	ä	92.9
	Score	14	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143
	Result No.	1	7	ю	4	S	9	7	æ	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29

genome polyprotein polypeptide - hepa	genome polyprotein genome polyprotein genome polyprotein	genome polyprotein genome polyprotein genome polyprotein	genome polyprotein genome polyprotein	genome polyprotein genome polyprotein genome polyprotein	genome polyprotein genome polyprotein	genome polyprotein
JQ0881 PC2219	GNWYCJ GNWYCJ S18030	GNWVCH S40770 JC5620	JQ1303 GNWVJ8	S21330 S41356 S41359	S41358 S41351	841349
997				700	9 79	~
874 876	3010	3011 3011 3014	3033	88 108 114	114	115
92.0	0000 0000 0000	922.9	92.6	90.3 90.3 90.3	90.3 90.3	90.3
143	143	143 143	143	139	139	139
30	1 60 60 1	36 37	8 O C	444	4 4 .	<b>4</b>

# ALIGNMENTS

RESULT 1 A45573		Fil616-1862/Product: nonstructural protein NG4a #status predicted <n4a> Fil863-2013/Product: nonstructural protein NS4b #status predicted <n4b> Fi2014-3010/Product: nonstructural protein NS5 #status predicted <ns5> Query Match 93.5%; Score 144; DB 1; Length 3010; Best Local Similarity 92.9%; Pred. No. 1.1e-11; Matches 26: Construction 1: New Matche</ns5></n4b></n4a>	1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28	RESULT 2 S41353 genome polyprotein - hepatitis C virus (genotype 2, N2) (fragment) N:Contains: core protein N:Contains: core protein C; Species: hepatitis C virus A; Variety: genotype 2, N2 C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
--------------------	--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Gaps

ö

Indels

No. 6.2e-13;

```
Similarity 92.9%; Pred. No. 6.26
26; Conservative 1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C; Accession: 541348
                     Best Local
Matches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 7
S41341
                                                                                                                                                    ò
                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Substitute polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
N;Contains: core protein
N;Contains: hepatitis C virus
N;Variable Nay-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
N;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
N;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
N;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
N;Reference number: S41341
A;Accession: S41357
A;Access
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Genome polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: 54135
B;Van Doorn, L.J.; Kletr, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41341
A; Accession: S41353
A; Molecule type: genomic RNA
A; Residues: 1-108 <-VAN>
A; Cross-references: EMBL:229456
A; Experimental source: genotype 2, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: core protein; status predicted <-WAT>
F; 1-108/Product: core protein #status predicted <-WAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92.9%; Score 143; DB 2; Length 108; 92.9%; Pred. No. 6.2e-13; ive 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 143; DB 2; Length 10
Pred. No. 6.2e-13;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A) Experimental source: genotype 2, N4
C) Superfamily: hepatitis C virus genome polyprotein
C) Keywords: capsid protein; core protein; polyprotein
F)1-108/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Cross-references: EMBL: Z29458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity 92.5
les 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Reference number: $41341
A; Accession: $41355
A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local S
Matches 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2-1355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

Length 108;

DB 2;

Score 143;

95.98;

Query Match

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N5
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41371
R;van Doorn, L.J.; Rieter, G.E.M.; Brouwer, J.T.
R;van Doorn, L.J.; Rieter, G.E.M.; Brouwer, J.T.
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
                                                                                                                                                                                                   V.Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                    LIPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                  genome polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus (genotype 1, N1) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                       ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                    S
                                                                                                                                                                                                                                                                                                                                    Rivan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1
A;Reference number: S41341
A;Accession: S41348
A;Molecule type: genomic RNA
A;Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Experimental source: genotype 1, N6 C; Superfamily: hepatitis C virus genome polyprotein C; Keywords C Stepsid protein; core protein; polyprotein F; 1-108/Product: core protein *status predicted xMAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCOSS-references: EMBL:229474
A/Experimental source: genotype 5, N5
C/Superfamily: hepatitis C virus genome polyprotein
C/Superfamily: hepatitis C virus genome polyprotein
E/1-112/Product: core protein #status predicted <AMJ>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 92.9%; Score 143; DB 2; Best Local Similarity 92.9%; Pred. No. 6.2e-13; Matches 26; Conservative 1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 143; DB 2;
Pred. No. 6.4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Mismatches
1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                           92.98;
92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Cross-references: EMBL: 229451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
Best Local Similarity 92.5.
Best Local Similarity 92.5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: S41371
A; Molecule type: genomic RNA
A; Residues: 1-112 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N;Contains: core protein
C;Species: hepatitis C virus
```

```
C; Keywords: capsid protein; core protein; polyprotein F;1-114/Product: core protein #status predicted <MAT>
                                                                           Query Match
Best Local Similarity
Matches 26; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Reference number:
A; Accession: S41368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Accession: S41342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                              genome
                                                                                                                                                                                                                                                                                                                                                            S41368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A41370

Genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
Noontains: core protein
Noontains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: 541370
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Reference number: 541341
A;Accession: 541370
A;Accession: 541370
A;Residues: 1-114 <VAN>
A;Residues: 1-114 <VAN>
A;Residues: 1-114 <VAN>
A;Cross-references: EMBL:229473; NID:9443908; PIDN:CAA82611.1; PID:9443909
A;Experimental source: genotype 5, N4
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <MAT>
A;Variety: genotype 1, N1
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41341
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
S;Van Doorn, L.J.; Kleter, G.E.M.; NID:9443850; PIDN:CAA82582.1; PID:9443851
A;Restdues: 1-112 < VANN
A;Restdues: 1-112 < VANN
A;Restdues: Genotypel, NI
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; Polyprotein
C;Keywords: capsid protein; status predicted <WAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 9
S41369
genome polyprotein - hepatitis C virus (genotype 5, N3) (fragment)
N:Contains: core protein
C:Species: hepatitis C virus
A:Variety: genotype 5, N3
C:Species: hepatitis C virus
A:Variety: genotype 5, N3
C:Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41369
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41349
A;Rolecule type: genomic RNA
A;Resious: 1-114 <VANN
A;Resious: 1-114 <VANN
A;Resious: genomic RNA
A;Resperimental source: genotype 5, N3
C;Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 143; DB 2;
Pred. No. 6.4e-13;
1; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 143; DB 2;
Pred. No. 6.5e-13;
1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                          92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 92.9 (atches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 26; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                   ö
                                                                                                                                                                                                                                                                                Wicontains: core protein:
C:Species: hepatitis C virus
A;Variaty; genotype 5, N2
A;Variaty; genotype 5, N2
C:Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41368
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A, Cross-references: EMBL:229445; NID:g443852; PIDN:CAA82583.1; PID:g443853
A, Experimental source; genotype 1, N10
C, Superfamily: hepatitis C virus genome polyprotein
C, Keywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein #status predicted <NAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A) Residues: 1-114 <-VAN>
A) Cross-references: EMBL: Z29471; NID: 9443904; PIDN: CAA82609.1; PID: 9443905
A) Experimental source: genotype 5, N2
C) Superfamily: hepatilis C virus genome polyprotein
C) Superfamily: core protein; core protein; polyprotein
F)1-114/Product: core protein #status predicted <-VAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C virus (genotype 1, N10) (fragment)
                                                                                                                                                                                                                                                                    polyprotein - hepatitis C virus (genotype 5, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C.Accession: S41342
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by A;Reference number: S41341
  Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; DB 2;
6.6e-13;
Score 143; DB 2;
Pred. No. 6.5e-13;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 143; DB 2;
Pred. No. 6.5e-13;
1; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatitis C virus (genotype 1, N;Contains: core protein S,Species: hepatitis C virus A;Vaziety: genotype 1, N10 C;Date: 19-May-1994 #sequence_revision 26-Jul-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 143;
                                                                                               28
                                                                                                                        28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.
                                                                                               1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92.9%;
92.9%;
92.9%;
1larity 92.9%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: genomic RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 26; Conserv
```

```
Search completed: August 7, 2003, 11:21:48
Job time : 9.54545 secs
                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŝ
              Wicontains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, N2
C; Species: hepatitis C virus
A; Variety: genotype 1, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S4134
A; Parled to the EMBL Data Library, January 1994
A; Rescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S4134
A; Residues: 1-115 < VAN
A; Experimental source: genouppe 1, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted < WANT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis C virus (genotype 1, NB) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, NB
C;Date: 19-Way-1594 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41350
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis C virus (genotype 2, N3) (fragment)
N;Contains: core protein
N;Contains: C virus
A;Variets: hepatitis C virus
A;Variety: genotype 2, N3
C;Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Cross-references: EMBL: 229453; NID: 9443868; PIDN: CAA82591.1; PID: 9443869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C. Accession: $41354
R.van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPa
A; Reference number: $41341
A.Accession: $41354
A; Molecule type: genomic RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö

    hepatitis C virus (genotype 1, N2) (fragment)

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.9%; Score 143; DB 2; Length 115, 92.9%; Pred. No. 6.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A:Experimental source: genotype 1, NB
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
'-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 143; DB 2;
Pred. No. 6.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity 92.9 ies 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Accession: S41350
A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Juery Match
Best Local Similarity
  polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, N3
C; Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Reference number: $41341
A; Accession: $41345
A; Accession: $41345
A; Accession: $41345
A; Molecule type: genomic RNA
A; Residues: 1-115 < VANA
A; Cross-references: EMBL: 229448; NID: 9443858; PIDN: CAA82586.1; PID: 9443859
A; Experimental source: genocitye 1, N3
A; Cross-references: EMBL: 229448; NID: 9443859
A; Residues: 1-115 < VANABL: 229448; NID: 9443859
A; Cross-references: EMBL: 229488; NID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859
A; Cross-references: EMBL: 229488; NID: 9443859
A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LIPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein - hepatitis C virus (genotype 1, N3) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 by
                                                                                                                                                                                                                                                                                                                                                  Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R) van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to
                                              A; Experimental source: genotype 2, N3
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                            Score 143; DB 2;
Pred. No. 6.6e-13;
1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
92.9%; Score 143; DB 2;
Best Local Similarity 92.9%; Pred. No. 6.6e-13;
Matches 26; Conservative 1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                       Ouery Match
Best Local Similarity 92.9%;
Matches 26; Conservative
A; Cross-references: EMBL: 229457
```

```
August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                          US-09-491-146A-28
154
1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                    OM protein - protein search, using sw model
                                                                                                                                                                                                            Title:
Perfect score: |
Sequence: |
                                                                                                                          Run on:
```

Scoring table:

BLOSUM62 Gapop 10.0 , Gapext 0.5

127863 tal number of hits satisfying chosen parameters: 127863 seqs, 47026705 residues Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

	40.0000	nor driver	σ	hepatitis					ㅁ	ч	h genome	c	h genome	h genome	h genome	herpes si		nicot	fugn	homo	spina		P03416 murine coro	P18448 murine coro		Q28193 bos taurus	P21573 xenopus lae	P18761 mus musculu	Q09459 caenorhabdi	P35824 bacillus ci	155 manduca	2362	grb	1816 homo	P26285 bos taurus
SUMMAKIES	£		POLG_HCVJT	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVTW	POLG_HCV1	IE63_HSV2H	FURI_HUMAN	RK4_TOBAC	RS7_FUGRU	RS7_HUMAN	RK4_SPIOL	NCAP_CVM3	NCAP_CVMA5	NCAP_CVMS	FURI_MOUSE	FURI_BOVIN	YB1_XENLA	CAH6_MOUSE	YQ38_CAEEL	SLAP_BACCI	RS7_MANSE	RS7_XENLA	RIBB_VIBPA	∾	F262_BOVIN
	ď	3 :	н	-	-	-	-	7	-	-	-	٦	7	<del>, -1</del>	~	~4	٦		-1	-	-	Н	_		ч	-4	~	-	-		Η.	-	-		-
	Length		3010	513	520	520	737	737	3010	3010	3011	3033	3033	3010	3011	512	794	282	194	194	293	424	454	454	793	797	303	317	1531	1616	190	194	218	1311	230
æ	Query	4			•	٠			95.9	٠			٠	٠	٠	38.3	•	34.4	32.5		٠		•		•	•				31.2	٠.	ö	•		29.9
	arone er		144	143	143			143	143	143	143	143	143	136	135	59	52	23	20	20	20	20	20	20	20	4		48	48	48	47	47	47	47	46
	Result			7	e	4	Ŋ	9	7	<b>∞</b> `	Φ.	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	53	30	31	32	33

O15530 homo sapien	Q922a0 mus musculu	O55173 rattus norv	P23377 rattus norv	014738 h serine/th	P78317 homo sapien	Q03717 mus musculu	095049 homo sapien	Q9r244 mus musculu	028653 o serine/th	P26786 saccharomyc	
PDPK_HUMAN	PDPK_MOUSE	PDPK_RAT	FURI_RAT	2A5D_HUMAN	RNF4_HUMAN	KCB1_MOUSE	ZO3_HUMAN	TRP2_MOUSE	2A5D_RABIT	RS7A_YEAST	RS7B_YEAST
7	٦	-		-	7	-	-	н	_	Н	-
556	559	559	793	602	190	857	933	1172	586	189	190
29.9	29.9	29.9	29.9	29.5	29.3	29.5	29.5	29.5	28.9	28.6	28.6
46	46	46	46	45.5	45	45	45	45	44.5	44	44

# ALIGNMENTS

셤

```
CELLULAR AMINOPEPTIDAGE.
CAPSID PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NON-STRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTENSE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
POTENTIAL.
                                                                                                                                                                                                                                                            Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 144; DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NKED (GLCNAC. ) (POT 94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                   InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_Bsvir.
Pfam; PF01543; HCV_capsid; 1
Pfam; PF01542; HCV_core; 1.
Pfam; PF01550; HCV_NSI; 1.
Pfam; PF01560; HCV_NSI; 1.
HCV_capsid
HCV_core.
                  IPR002519; HCV_env.
IPR002531; HCV_NS1.
IPR002518; HCV_NS2.
IPR0004109; HCV_NS3.
IPR001409; HCV_NS48.
IPR001409; HCV_NS48.
                                                                                                                                                                Pfam; PF01538; HCV_NS2; 1. Pfam; PF02907; HCV_NS3; 1. Pfam; PF01006; HCV_NS4a; 1. Pfam; PF011001; HCV_NS4b; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF00598; Viral_RGRP; 1. ProDom; PF086062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2788
326573 M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 93.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 93.5
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                         729
1006
1615
1862
2013
369
369
1107
11165
1319
196
209
                                                                                                                                                                                                                                                    SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ä
                                                                                                                                                                                                                                                                                           3D-structure.
INIT_MET
                            InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                    CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseélsb-sib.ch).
                       01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope 41ycoprotein El (GP32) (GP35); Envelope 91ycoprotein E2
Hepatitis C virus (isolate HC-J2) (HCV).
Hepatitis C virus (isolate HC-J2) (HCV).
Hepatitis C virus (isolate HC-J2) (HCV).
                                                                                                                                                                                                                                                                                                   genotypes.";
Virology 188:331-341(1992).
-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPMOREIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTEWIAL).
MATRIX PROTEIN (FOTEWIAL).
MAJOR ENVELOPE PROTEIN E (POTEWIAL).
NONSTRUCTURAL PROTEIN NS1 (POTEWIAL).
                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kural K., Okada S.I., Yamamoto K., Lizuka H.,
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
"Full-length sequence of a hepatitis C virus genome having poor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D10074; BAA00968.1; ...

R InterPro; IPR002521; HCV_copsid.

InterPro; IPR002521; HCV_core.

R InterPro; IPR002519; HCV_core.

R InterPro; IPR002519; HCV_core.

R InterPro; IPR002519; HCV_core.

R Pfam; PF01543; HCV_core; 1.

R Pfam; PF01560; HCV_core; 1.

R Pfam; PF01560; HCV_core; 1.

R Probom; PD186062; HCV_NS1; 1.

R Proportein; Glycoprotein; Cat protein; Envelope protein; Transmembrane; Nonstructural protein.

I INIT_MET 1.

R PEMOVED FROM CAPSID PROTEIN C BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
   Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                          NCBI_TaxID-11111;
POLG_HCVJ2
P27959;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
```

ö

Gaps

; 0

le-12;

Pred. No.

1; Mismatches

ò g SEPTER S

```
NCBI_TaxID=31644;
                                                                                                                                                                                   POLG_HCVHK
Q01403;
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                      Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                          RESULT 4
POLG_HCVHK
                                                                Matches
                                                                                                                  g
                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lab-sib.ch).
                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REMOVED FROM CAPSID PROTEIN,
REMOVED FROM CAPSID PROTEIN C BY THE
CELLICARA MAINOPERTIDASE.
MATRIX PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MADON ENVELOPE PROTEIN E (POTENTIAL).
BY SIMILARITY.
                                                    Gaps
                                                                                                                                                                                              01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment).
                                                                                                                                                                                                                                                                                                                                         MEDLINE-93019030; PubMed-1383400;
ADE K., Inchauspe G., Fujisawa K.;
ADE K., Inchauspe G., Fujisawa K.;
AGEOMIC characterization and mutation rate of hepatitis C virus
'Genomic characterization and mutation and round on the contracted hepatitis and epidemic
of non-A, non-B hepatitis in Japan.";
Of non-A, non-B hepatitis in Japan.";
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN WAND GLYCOPROTEIN E. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A AND MINA.
                                                                                                                                                                                                                                                                           Hepatítis C virus (isolate HCV-476) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                    ;
0
                          Length 513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                    1; Indels
55704 MW; 943F31E3514CDEF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC. . .
                       Score 143; DB 1;
Pred. No. 2.1e-13;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                      520 AA
                                                                                           PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                           1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D10688; BAA01530.1; --
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_MS1.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                        92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nonstructural
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF01539; HCV_env; 1
                                                    Conservative
                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    513 AA;
                        Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Transmembrane;
                                                                                                                                                                                                                                                                                                       Hepacivirus
                                                                                                                                                                    POLG_HCVH4
Q01404;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRANSMEM
                                                                                                                                           RESULT 3
POLG_HCVH4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
S
```

DRAKKE BETTA E ETT E

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation for European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CAPSID MATRIX PROTEIN ( POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELODE PROTEIN B (POTENTIAL).
BY SIMILARITY.
N-LINKED (GLCNAC. . . ) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                             01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66-One polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GF92) (GF95); Envelope glycoprotein E2 (GF68) (GF70) (NS1)] (Fragment)
Hepatitis C virus (isolate HCV-KF) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-93019030; PubMed-1383400;
Abe K., Inchauspe G., Fujisawa K.;
"Genomic characterization and mutation rate of hepatitis C virus
"Genomic characterization and mutation of non-A, non-B hepatitis in Japan.";
J. Gen. Virol, 73:275-2729(1992).
-I-SUBNINT: THE VIRIO PROFED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROFEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             protein. REMOVED FROM CAFSID PROTEIN C BY THE
                                                                                                                                             ö
                                                                                  Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Coat protein; Envelope protein;
                                                                                                                                             Indels
520 520
520 AA; 56499 MW; AAl35246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE
                                                                                  Score 143; DB 1;
Pred. No. 2.1e-13;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                     520 AA
                                                                                                                                                                                                                                5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                               1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF01543; HCV_capsid; 1. Pfam; PF01542; HCV_core; 1. Pfam; PF01539; HCV_env; 1. Pfam; PF01569; HCV_NS1; 1. ProDom; PD186062; HCV_NS1; 1. Polyprotein; Glycoprotein; Coat Transmembrane; Nonstructural pr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D10687; BAA01529.1; -. PIR; JQ1925; JQ1925.
                                                                                     92.98;
                                                                                                                                          26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               115
191
383
369
196
233
233
418
                                                                                                                Similarity
```

us-09-491-146a-28.rsp

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a licensee agreement (see http://www.isb-sib.ch/announce/or send an email to licenseelab-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/P2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                          01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66-nome polyprotein (Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NSI and NS2] (Fragment).
Hepatitis C virus (isolate H-J5) (HCV).
                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROTEIN AND GLYCOPROTEIN ENVELOPE CONPINENCE OF TWO TOTAL DISCLAIM OF THE WINGTON: THE WINGTON NS2A, NS2B, NS2B, NS4B AND NS4B ARE HYDROPHORIC, SUGGESTING A POSSTBLE MEMBRARE FELATED FUNCTION.

1. SUBUNIT: THE WIRINN OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-92230232; PubMed-1314459; Okamento K., Lizuka H., Okameto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Chamaka T., Fukuda S., Tsuda F., Mishiro S.; Fransa T., Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ΒY
                                                                                                     Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Coat protein; Envelope protein;
                                                                                                                                         1; Indels
                                                                     56476 MW; 1D2BD0A6FF27349B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CELLULAR AMINOPEPTIDASE
                                                                                                       Score 143; DB 1;
Pred. No. 2.1e-13;
                                                                                                                                                                                                                                                                                             737 AA
                                                                                                                                       1; Mismatches
                                                                                                                                                                       1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                           protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002522; HCV_capsid.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Co
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nonstructural
                                                                                                       92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; D10075; BAA00969.1; -
                                                                                            Query Match
Best Local Similarity 92.99
                                                                                                                                                                                                                                                                                             STANDARD;
 424
431
449
520
520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             116
192
384
734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                             POLG_HCVJ5
                                CARBOHYD
NON_TER
SEQUENCE
 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INIT_MET
                                                                                                                                                                                                                                                           RESULT 5
POLG_HCVJ5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                           385
   FF F F S
                                                                                                                                                                                                       셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-92230232; PubMed-1314459;
AMEDLINE-92230232; PubMed-1314459;
AMEDLINE-92230232; PubMed-1314459;
AMEDLINE-92230232; PubMed-1314459;
ATANARA T., Fukuda S., Tsuda S., Tsuda F., Mishiro S.;
Tranaka T., Fukuda S., Tsuda F., Mishiro S.;
Tranaka T., Fukuda S., Tsuda F., Mishiro S.;
Tranaka T., Fukuda S., Tsuda S., Tsuda B., Mishiro S.;
Tranaka T., Fukuda S., Tsuda S., Tsuda B., Mishiro S.;
Tranaka T., Fukuda S., Tsuda S., Mishiro S.;
Tranaka T., Fukuda S., Tsuda S., Mishiro S.;
Tranaka T., Fukuda S., Tsuda S., Mishiro S.;
Tranaka T., Fukuda S., Tranaka T., Yamamoto S.;
Tranaka T., Fukuda S., Tranaka B., Mishiro S.;
Tranaka T., Fukuda S., Tranaka T., Yamamoto S.;
Tranaka T., Fukuda S., Tsuda S., Mishiro S., Mishiro S.;
Tranaka T., Fukuda S., Tranaka T., Yamamoto S.;
Tranaka T., Fukuda S., Tsuda S., Mishiro S., 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-07-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                  (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                   (POTENTIAL)
                                                                                                                                                                                                               (POTENTIAL)
                                              ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                         Length 737;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                   3AF699D82AD501B1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                       Score 143; DB 1; Lv
Pred. No. 3.1e-13;
1; Mismatches 1;
                                                                                                                               (GLCNAC.
                                                                                                                                                                                                               (GLCNAC.
                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                              (GLCNAC
                      (GLCNAC
                                                                                       (GLCNAC
                                                                                                           (GLCNAC
                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               737 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  l and NS2] (Fragment).
virus (isolate HC-J7) (HCV).
                                                                                                      N-LINKED
                                                              N-LINKED
N-LINKED
                                          N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_cnv.
InterPro; IPR002531; HCV_NSI.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                   ¥.
                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                                                                                                                 81207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus.
NCBI_TaxID=11114;
                                                                                                                                                                                                                                                                                                                                             737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVJ7
P27961;
                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                      CARBOHYD
NON_TER
                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                    CARBOHYD
                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCVJ7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                       CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
MEDINE-91140689; PubMed-1847440;
MEDINE-9140689; PubMed-1847440;
Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J.,
Onishi E., Andoh T., Yoshida I., Okayama H.;
"Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                            THE
                                                                                                                                                                                                                             (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                              ( POTENTIAL
                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1487-1500.
MEDLINE-9623224; PubMed-8647104;
BOCTOWS! P., Heiland M., Ochlmann K., Becker B., Korneteky L.;
*Non-structural protein 3 of hepatitis C virus inhibits
                                                                                                                                                   POTENTIAL
                                                                                                                                                          (POTENTIAL
                                                                                                                                                                                                                                                                      (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                  ö
                                                            ΒŸ
                                                                                                                                                                                                                                                                                                                               Score 143; DB 1; Length 737;
Pred. No. 3.1e-13;
1; Mismatches 1; Indels
                                                Nonstructural protein.
1 REMOVED FROM CAPSID PROTEIN C
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Probom; PF01560; HCV_NS1; 1.
Probom; P0186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                                            67DFAE11854122F2 CRC64;
                                                                    CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                          GLCNAC
                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                         Pred. No. 3.16
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                     1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                     N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                POTENTIAL.
                                                                                                                                                                                                                                                  -LINKED
                                                                                                                                                                                                                                                                                                             81691 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                             92.9%;
                                                                                                                                                                                                                                                                                                                                        1 Similarity 92.9
26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                        649
737
737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-11105;
                                                 Transmembrane;
INIT_MET 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCVBK
                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                     CHAIN
TRANSMEM
                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                  NON_TER
SEQUENCE
                                                                                                                                                  CARBOHYD
                                                                                                                                                           CARBOHYD
                                                                                                                                                                      CARBOHYD
                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                          Local
                                                                                                                                                                                                                                                                                                                                                                                                                                               P26663;
                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 7
POLG_HCVBK
                                                                               CHAIN
                                                                                                            CHAIN
                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                 ARANATATATATATATATA
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                         525E
```

```
This. SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                       X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.

MEDLINE-98227846; PubMed=9568891;
A Yan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,
A Steinkuehler C., Tomel L., de Francesco R., Kuo L.C., Chen Z.;
"Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
virus: a 2.2-A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-847(1998).

- 1- FUNCTION: THE SMALL PROTEINS NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAX A ROLE IN THE VIRAL RNA REPLICATION.
C -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral costition, Cys or Thr in Pl and Ser or Ala in Pl'.
C -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                     {RNA}(N).
-!- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                              MEDLINE-97015088; Pubmed-8865916; Love R.A., Parge H.E., Wickersham J.A., Hostcmsky Z., Habuka N., Moomaw E.W., Adachi T., Hostcmska Z.; The crystal structure of hepatitis C virus NS3 proteinase reveals trypsin-like fold and a structural zinc binding site."; Cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -! - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                  K-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR001490; HCV_NS4b.
IPR002868; HCV_NS5a.
IPR002166; HCV_RdRP.
IPR007095; RNA_POl_DS_PS.
IPR007094; RNA_POl_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF01543; HCV_caps1d; 1.
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1A10; 25-MAR-98.
1UXP; 14-JAN-98.
1NS3; 08-APR-98.
1C2P; 15-NOV-00.
1CSJ; 08-NOV-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-JUN-00,
20-APR-99,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09-APR-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10-APR-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                $29.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1GX5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1007;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1GX6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BOHM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              nterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PDB;
```

Length 3010;

```
1187 1188
1189 1197
1198 1202
1203 1204
1680 1688
3010 AA; 327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                            Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus.
                                                     STRAND
                STRAND
                                                                                                                                                                                                                 RESULT 8
POLG_HCVJA
    TURN
                                            TURN
                                                                                                                                                                                                                                           FT
FT
FT
SO
                                                                                                                                                                        g
                                                                                                                                                  à
                                                                                                                              CELLULAR ANINOPERTIDASE.
CAPSID PROTEIN ( POTENTIAL).
MATRIX PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/S2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA DIRECTED RNA POLYMERASE (POTENTIAL).
                                                               Polyprotein; Giycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                      REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENT
                                                                                                                                                                                                                                                                           CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LINKED
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00999; Viral_RGRP; 1.
ProCom; PD186062; HCV_N31; 1.
SWART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1085
1087
1092
1094
                                                                                                                                                                                                                                                                                                                                                                                                                         4430
4430
546
546
576
653
663
674
674
674
674
688
688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1031
10030
10050
10050
10050
10075
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
10090
1
                                                                                                          3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                    ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
STRAND
                                                                                                                                                                                                                                                                TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TURN
STRAND
HELIX
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
STRAND
HELIX
TURN
                                                                                                                                             CHAIN
CHAIN
CHAIN
                                                                                                                                                                                    CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HELIX
                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TURN
```

```
DOUGLECTORY STANDARD; PRT; 3010 AA.

POLG_HCVJA STANDARD; PRT; 3010 AA.

To 1-40G-1992 (Rel. 23, Created)

To 1-40G-1992 (Rel. 23, Last sequence update)

To 1-40G-1992 (Rel. 23, Last sequence update)

To 1-40G-1992 (Rel. 24) Last annotation update)

To 28-FEB-2003 (Rel. 41, Last annotation update)

To 28-FEB-2003 (Rel. 41, Last annotation update)

To 28-FEB-1003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISCUSSION OF SEQUENCE.

MEDLINE-91192160; PubMed-1849488;

Mato N., Hillhata M., Nakagawa M., Ootsuyama Y., Muraiso K.,

Ohkoshi S., Shimotchno K.,

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"FEBS Left. 280:325-328(1991).

"FEB
                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [RNA](N).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSIGNS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-91088550; PubMed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S., Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
Score 143; DB 1;
Pred. No. 1.4e-12;
1; Mismatches 1
                                                                                                                                                                                                                                                           1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                      EMBL; D90208; BAA14233.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; A39253; GNWVCJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11116;
```

```
PKPQRKTKRNAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RNA) (N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus
                                                                                                                                                                                                                      POLG_HCVH
                               -
                                                                                 ນ
                                                                                                                                                               POLG_HCVH
                             ò
                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLYPOTECHI, Glycoprocein; Transferase; RNA-directed RNA polymerase;
W Core protein; Goat protein; Envelope protein; Helicase; APP-binding;
Transmembrane; Nonstructural protein C By THE
CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).
TCHAIN 192 383 MAJOR ENVELOPE PROTEIN E, POTENTIAL).
TCHAIN 193 100 NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
TCHAIN 1615 PROTESSE/HELICASE NS3 (POTENTIAL).
TCHAIN 1615 NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
TCHAIN 1663 2013 NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
TCHAIN 1663 2013 RNA-DIRECTED NNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ( POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 143; DB 1; Length 3010;
Pred. No. 1.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                Interpro; IPR007095; RNA_pol_DS_PS.
Interpro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1
                                                                                                                                  HCV_core.
HCV_env.
HCV_NS1.
HCV_NS3.
HCV_NS3.
HCV_NS3.
HCV_NS4.
HCV_NS4.
HCV_NS4.
                                                                                                           HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00271; hellcase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF01542; HCV_core; 1.
PF01539; HCV_env; 1.
PF01550; HCV_NS2; 1.
PF02907; HCV_NS2; 1.
PF01006; HCV_NS4; 1.
PF01001; HCV_NS4; 1.
PF01001; HCV_NS4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2041
2077
2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2013
3010
369
1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00487; DEXDC;
                                                                              InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                           IPR001650;
                                                                                                                                                                                                                InterPro; IPR002518;
InterPro; IPR004109;
                                                                                                                                                                                                                                                                          InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                       IPR001490;
                                                                                                                                                                                                                                                                                                                                 IPR002868;
                                                                                                                                                                                                                                                                                                                                                            IPR002166;
                                                                                                                                                                  PR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ¥
                       MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3010
                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
CORRECT OF THE FET THE
```

Gaps

ö

1; Indels

1; Mismatches

Conservative

Best Local Similarity Matches 26; Conserv

Query Match

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sits content is in no way
Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on it use by non-profit institutions as long as its content is in no wa modified and this statement is no removed. Usage by and for commercial
                                                                                                                                                                                                                                                       01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annoctation update)
66-00me polyprotein (Contains: Capsid protein C (Core protein) (P22);
67-009 (GP68) (GP70) (NS1); Protein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66); Nonstructural protein NS5A (P66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong 2., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             oligonucleotide: the crystal structure provides insights into the mode of unwinding.";
Structure 6:89-100(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -!- FUNCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-I- FUNCTION: PROTEASAE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS3-NS4A, NS4A-NS4A, NS4A-NS4B, NS4A-NS5A AND NS5A-NS5A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-9915431; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPMOTEIN ENVELOPE. OTHE SEVELEDPE CONSISTS OF TWO FROTEINS: EL AND E2. THE NUCLESCARSID IS A COMPLEX OF PROTEIN C AND MENA. PTM: THE STRUCTURAL PROTEINS C, EL AND E2 ARE PRODUCED BY PROTECLYIC PROCESSING BY THE MOST SIGNAL PPEPIDASES. SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY 039. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Prince A.M.;
"Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates.";
Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
MEDLINE-92052256; Pubmed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                              3011 AA
58
                                                      32
                            PRT;
                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACTIVATION OF NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-11108;
```

us-09-491-146a-28.rsp

```
. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 143; DB 1; L
Pred. No. 1.4e-12;
1; Mismatches 1;
    N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
    623

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6242

6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1594 159
1598 159
1606 161
1614 161
1622 162
1627 162
1635 163
1640 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
POLG_HCVJ6
ID POLG_HCVJ6
CARBOHYD
STRAND
TURN
HELIX
TURN
TURN
STRAND
TURN
STRAND
STRAND
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
HELIX
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
HELIX
TURN
STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
HELIX
SEQUENCE
    (See http://www.isb-sib.ch/announce/
                                         EMBL: M67463 AAA45534.1; -.

DR PIB; A56814; GNWYCH.

DR PDB; HAIR; 125-NO-98.

DR PDB; HAIR; 17-JUN-98.

DR PDB; HAIR; 17-JUN-98.

DR MEROPS; 292-001; -.

DR MEROPS; 292-001; -.

DR INCEPPO: IPR0015212; HCV_core.

DR INCEPPO: IPR0015213; HCV_core.

DR INCEPPO: IPR0015213; HCV_core.

DR INCEPPO: IPR0015219; HCV_core.

DR INCEPPO: IPR0015219; HCV_core.

DR INCEPPO: IPR001519; HCV_LOST.

DR INCEPPO: IPR001519; HCV_LOST.

DR INCEPPO: IPR001519; HCV_LOST.

DR INCEPPO: IPR001516; HCV_LOST.

DR INCEPPO: IPR00156; HCV_LOST.

DR Ffam; PF0154; HCV_Core.

DR Ffam; PF0154; HCV_CORP.

DR Ffam; PF0154; HCV_LOST.

DR Ffam; PF0155; HCV_LOST.

DR Ffam; PF0156; HCV_NS3; I.

DR Ffam; PF01001; HCV_NS3; I.

DR Ffam; PF01001; HCV_NS4; I.

DR Ffam; PF01005; HCV_NS4; I.

DR Ffam; PF010098; HCV_NS4; I.

DR Ffam; PF010098
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CELLULAR ANINOPERTIDASE.

CARSID PROTEIN C.

ENVELOPE GLYCOPROTEIN E1.

ENVELOPE GLYCOPROTEIN E2.

PROTEIN P7.

NONSTRUCTURAL PROTEIN NS2.

PROTENTIAL.

CHARGE RELAY SYSTEM (BY SIMILARITY).

N-LINKED (GLCNAC...) (POTENTIAL).

N-LINKED (GLCNAC...) (POTENTIAL).
ities requires a license agreement (Se send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
CARBOHYD
CARBOH
    entities
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
CHAIN
CHAIN
```

ö

Gaps

```
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus.
NCBI_TaxID=11115;
                                                                                                                                                                                                                                                1087
1111
1169
1234
1320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                             ACT_SITE
                                                                                                                                                                                                                                                                         SITE
                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                 ACT_SITE
                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                      TRANSMEN
                                                                                                                                                                                                                                                                              BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVJ8
   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                        MEDLINE-9204440; PubMed-1658196;
A MEDLINE-9204440; PubMed-1658196;
A Machida A., Miyakawa Y., Mayumi M.,
Machida A., Miyakawa Y., Mayumi M.,
Machida A., Miyakawa Y., Mayumi M.,
Machida B., Miyakawa Y., Mayumi M.,
Machida Sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";
J. Gen. Virol. 72:2697-2704(1991).
J. Gen. Virol. 72:2697-2704(1991).
HYDROPHOBIC, SUGGESTING A POSTBILE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROPE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROPE IN THE VIRAL RNA REPLICATION.
C -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral postition, Cys or Thr in Pl and Set or Ala in Pl'.
C -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
          01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 23, Last sequence update)
01-60G-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein EI (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22..); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS58 (P65) (P77); Nonstructural protein NS58 (P65) (P77); Nonstructural protein NS58 (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-36) (HCV).

    -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ro; IPR007095; RNA_pol_DS_PS.
ro; IPR007094; RNA_pol_PSvir.
PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_RGRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_core.
HCV_core.
HCV_env.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D00944; BAA00792.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR004109; H. InterPro; IPR000745; HC InterPro; IPR00145; HC InterPro; IPR001465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  interPro; IPR001410;
interPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                        NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF01542;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01001;
                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
```

```
D POLG_HCVJ8 STANDARD; PRT; 3033 AA.

POLG_HCVJ8 STANDARD; PRT; 3033 AA.

10 26661;

T0 1-40G-1992 (Rel. 23, Last sequence update)

T0 1-40G-1992 (Rel. 23, Last sequence update)

T0 1-40G-1992 (Rel. 24, Last annotation update)

T0 1-40G-1992 (Rel. 24, Last annotation update)

Envelope glycoprotein [Contains: Capsid protein C (Core protein) (P22);

Envelope glycoprotein E2

Envelope glycoprotein E2

Envelope glycoprotein E2

EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)

EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)

EC 3.4.22.-); Nonstructural protein NS5A (P56); Nonstructural protein

ENS5B (P65) (P70) (RNA directed RNA polymerase) (EC 2.7.7.48);

Hepatitis C virus (isolate HC-J8) (HCV).

O Viruses; SRNA positive-strand viruses, no DNA stage; Flaviviridae;
TOTAL SECTION 1. Transferase; RNA-directed RNA polymerase; SMART; SMO0487; DEXDC. 1.

W Polyprotein; Glycoprotein Transferase; RNA-directed RNA polymerase; Core protein; Goat protein; Envelope protein; Helicase; AFP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

T INIT_MET 1 CHAIN 116 CAPSID PROTEIN C BY THE CELLURA RAINOPEPTIDASE.

T CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).

T CHAIN 384 733 MANOR ENVELORE PROTEIN SI (POTENTIAL).

T CHAIN 1011 1619 RNOTRUCTURAL PROTEIN NSI (POTENTIAL).

T CHAIN 1011 1619 RNOTRUCTURAL PROTEIN NSI (POTENTIAL).

T CHAIN 1619 RNOTRUCTURAL PROTEIN NSI (POTENTIAL).

T CHAIN 1619 RNOTRUCTURAL PROTEIN NSI (POTENTIAL).

T CHAIN 1670 1867 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).

T CHAIN 1867 2017 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).

T CHAIN 2018 S033 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 143; DB 1;
Pred. No. 1.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. 1.46
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 92.55,
Best Local Similarity 92.55,
Conservative
```

SIMILARITY). SIMILARITY). SIMILARITY).

(BY (BY (BY

(POTENTIAL)

GLCNAC GLCNAC GLCNAC POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL)

(POTENTIAL)

(POTENTIAL)

(POTENTIAL) POTENTIAL (POTENTIAL) POTENTIAL)

```
01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
66-00me polyprotein (Contains notation update)
66-00me polyprotein (Contains Capsid protein C (Core protein) (P22);
67-00me polyprotein El (GP32) (GP35); Envelope glycoprotein E2
67-00me polyprotein El (GP32) (GP35); Envelope glycoprotein E2
67-00me polyprotein El (GP32) (GP35); Envelope glycoprotein E2
68-3-4-21-98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P57); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)}.
Hepatitis C virus (Isolate Taiwan) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1A173E7E3381FD1A CRC64:
                                          CHARGE RELAY SYSTEM (
CHARGE RELAY SYSTEM (
CHARGE RELAY SYSTEM (
ATP (POTENTIAL)
DECH BOX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 143; DB 1;
Pred. No. 1.4e-12;
                                                                                                                                       N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. 1.4e
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-92230206; PubMed-1314449;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2811
330177 W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                649
1091
2038
2359
2359
231
2811
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVTW
P29846;
                                                                                ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                         CARBOHYD
                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 12
        ò
                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein; Givenor, 1.

Polyprotein; Givenorous, 2007.

Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE CELULAR AMINOPEPRIDASE.

CHAIN 116 191 MAPRIX PROTEIN (POTENTIAL).

CHAIN 192 383 MAJOR ENVELOPE PROTEIN E (POTENTIAL).

CHAIN 384 733 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).

CHAIN 734 1010 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).

CHAIN 1011 1619 PROTEASE NSI (POTENTIAL).

CHAIN 1011 1619 PROTEASE NSI (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                     SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.

SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Interpro; IPR002531; HCV_NS1.
Interpro; IPR002518; HCV_NS2.
Interpro; IPR004109; HCV_NS3.
Interpro; IPR001490; HCV_NS3.
Interpro; IPR001490; HCV_NS4a.
Interpro; IPR001866; HCV_NS5a.
Interpro; IPR002666; HCV_NS5a.
Interpro; IPR007095; RNA_POl_DS_PS.
Interpro; IPR007095; RNA_POl_DS_PS.
Pfam; PF01543; HCV_Capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_MS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; D10988; BAA01761.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF01542; HCV_core;
PF01539; HCV_env; 1
PF01560; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               733
1010
1619
1666
2017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PIR; A40250; GNWVJ8.
HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF02907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
DORREL CORRECT CORRECT
```

ö

Gaps

ö

Length 3033; 1; Indels

```
EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                D,R
                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CELLULAR AMINOPEPTIDASE.
CORE PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN (AD. 1922).
NONSTRUCTURAL PROTEIN NS.1.22 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.1.22 (POTENTIAL).
PROTEASE/HELICASE NS.3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.48 (POTENTIAL).
RNA-DIRECTED RNA POLLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P Fram, PF01542; HCV-Cure; 1.

R Pfam; PF01542; HCV-Cure; 1.

R Pfam; PF01560; HCV-N01; 1.

R Pfam; PF01560; HCV-N02; 1.

R Pfam; PF01006; HCV-N32; 1.

R Pfam; PF01006; HCV-N34; 1.

R Pfam; PF001006; HCV-N35; 1.

R Pfam; PF00271; HcV-N35; 1.

R Pfam; PF00271; HcV-N35; 1.

R Pfam; PF00998; Viral RdRP; 1.

R Pf00998; Viral RdRP; 1
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M84754; -; NOT_ANNOTATED_CDS.
PIR; A40244; GNWTW.
PDB; INN64; 25-FEB-03.
PDB; INN64; 25-FEB-03.
PDB; INN64; 25-FEB-03.
PDB; INN64; 25-PEB-03.
MENOPS; U39_001; -.
MENOPS; U39_001; -.
InterPro; IPR001410; DEAD.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00253; HCV_capsid.
InterPro; IPR00253; HCV_NS2.
InterPro; IPR00153; HCV_NS3.
InterPro; IPR00145; HCV_NS4.
InterPro; IPR001490; HCV_NS5.
InterPro; IPR001490; HCV_NS5.
InterPro; IPR001490; HCV_NS5.
InterPro; IPR001905; HCV_NS5.
InterPro; IPR001909; HCV_NS5.
InterPro; IPR001909; HCV_NS5.
InterPro; IPR001909; HCV_NS5.
InterPro; IPR001995; HCV_NS5.
InterPro; IPR001095; HCV_NSS.
INTERPRO; IPR001095; HCV_NSS.
INTERPRO; IPR001095; HCV_NSS.
INTERPRO; IPR001095; HCV_NSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1616
1863
2014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        THE REAL PROPERTY OF THE REAL
```

NKED (GLCNAC. NKED) (GLCNAC.	Score 136; DB 1; Length 3010; Pred. No. 1.5e-11; 1; Mismatches 1; Indels 0; Gaps 0;	eccolvg 28          	3.  4. Hevi Standard; PRT; 3011 AA.  6. Hevi Standard; PRT; 3011 AA.  6. Hevi Standard; Bat sequence update)  6. Hogical 1992 (Rel. 23, Created)  6. Hogical 1992 (Rel. 23, Created)  6. Hogical 2003 (Rel. 22, Last sequence update)  6. Hogical 2003 (Rel. 22, Lest sequence update)  6. Hogical 2003 (Hepacining 2003)  6. Hogical 2004 (Hepacining 2003)  6. Hogical 2004 (Hepacining 2003)  6. Hogical 2004 (Hepacining 2004)  6.	ROTEASE BELONGS TO PEPTIDASE FAMILY S29.  stropyright. It is produced through a collaboration struct of Bioinformatics and the EMBL outstation matics Institute. There are no restrictions on its institutions as long as its content is in no way reement is not removed. Usage by and for commercial icense agreement (See http://www.isb-sib.ch/announce/icense@isb-sib.ch)
	88.3%; Similarity 92.6%; 5; Conservative	KPQRKTKRNAHRRPQDVKFPGGGIVG            	13 V1 GG_HCV1 STANDARD; PRT; 3011 GG_664; -AUG-1992 (Rel. 23, Created) -AUG-1992 (Rel. 23, Last sequence upda -AUG-1992 (Rel. 23, Last sequence -AUG-1992 (Rel. 23, Lest s	Y: THE PROTE TENTY IS SWISS INSTIT BIOINFORMATI PROFIT INSTITUTES A ILCEN ANALSTORY IS ANALSTORY INSTITUTES A ILCEN ANALSTORY INSTIT
CARBOHYD	Query Match Best Local S1 Matches 25;	9 6 7 - 7	CC 13 2000013 2000013 2000013 2000013 2000013 2000013 2000013 2000013 2000013 2000013 2000013 2000013 2000013 2000013 2000013	This SWIZS-PR This SWISS-PR Detween the the European use by non- modified and entities requ or send an em
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	Ou Be: Mai	Oy Dp	E	888888888888

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FURI_HUMAN STANDARD; PRT; 794 AA.
P09958; Q14336;
01-MAF-1989 (Rel. 10, Created)
01-MAF-1999 (Rel. 14, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Furin precursor (EC 3.4.21.75) (Paired basic amino acid residue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Early protein; Transcription regulation; Activator; DNA-binding
SEQUENCE 512 AA; 54958 MW; 459651470A503BA7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ..
                                                                                              ö
                                                   Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    38.3%; Score 59; DB 1; Length 512; 40.6%; Pred. No. 0.47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12; Indels
        327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                          Herpes simplex virus (type 2 / strain HG52).
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
NCBI_TaxID-10315;
           MW; ... Score 135; DB 1; Leng
                                                                                                                                                                                                                                                                                                                         01-DEC-1992 (Rel. 24, Created)
01-DEC-1992 (Rel. 24, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Transcriptional regulator IE63 (VMW63) (ICP27).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNAH----RRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                        512 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3; Mismatches
                                                                                                                                                        1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D10471; BAA01269.1; -. EMBL; 286099; CAB06702.1; -.
                                            Query Match
Best Local Similarity 85.7%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 40.6
Matches 13; Conservative
                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR; JQ1498; WMBEXA.
        3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                 IE63_HSV2H
P28276;
      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                         RESULT 14
IE63_HSV2H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
FURI_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BAPPE
                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                            REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN NSI/F2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/F2 (POTENTIAL).
PROYEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA_DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                   InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_env.
InterPro; IPR002518; HCV_NS1.
InterPro; IPR00218; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR002166; HCV_NS5a.
InterPro; IPR002166; HCV_NS5a.
InterPro; IPR001699; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
                                                                                                         InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF02907; HCV_NS3; 1.
PF01006; HCV_NS4s; 1.
PF01001; HCV_NS4s; 1.
PF01506; HCV_NS5s; 1.
PF01505; HCV_NS5s; 1.
PF09091; Pt1al_RRFP; 1.
om; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF01543; HCV_capsid; Pfam; PF01542; HCV_core; 1. Pfam; PF01539; HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ProDom; PD186062; HCV_SMART; SM00487; DEXDC;
A39166; GNWVC3.
1A1V; 16-FEB-99.
1HEI; 25-NOV-98.
                                                        MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11616
11863
2014
11083
11083
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF01560;
PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
```

LANGURAR MANAGERAR AND CORRECT CORRECT

ä

EMBL outstation

and the

the Swiss Institute of Bioinformatics

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-92332543; PubMed=1629222;
Leduc R., Molloy S.S., Thorne B.A., Thomas G.;

an intraion of human lurin precursor processing endoprotease occurs by
an intramolecular autoproteolytic cleavage.";
J. Blol. Chem. 267:14304-14308(1992)

-!-FONCTION: FURIN IS LIKELY TO REPRESENT THE UBIQUITOUS ENDOPROTEASE
-!-FONCTION: FURIN IS LIKELY TO REPRESENT THE UBIQUITOUS ENDOPROTEASE
ACTIVITY WITHIN CONSTITUTIVE SECRETORY PATHWAYS AND CAPABLE OF
CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF.
-!-CATALYTIC ACTIVITY: Release of mature proteins from their
complement component C3 and von Willebrand factor from their
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBCELLULAR LOCATION: SEEMS TO BE LOCALIZED INTRACELLULARLY TO THE TRANS GOLGI NETWORK. PROPEPTIDE CLEAVAGE IS A PREREQUISITE FOR EXIT OF FURIN MOLECULES OUT OF THE ENDOPLASMIC RETCULUM (ER). SECOND CLEAVAGE IN THE PROPEPTIDE OCCUR IN THE TRANS GOLGI NETWORK (TGN), WHICH IS FOLLOWED BY THE RELEASE OF THE PROPEPTIDE BOUND TO FURIN AND THE ACTIVATION OF FURIN.
                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 296-794 FROM N.A. MEDLINE-87053858; PubMed-3023061; MEDLINE-87053858; PubMed-3023061; Roebroek A.J.M., Schalken J.A., Leunissen J.A.M., Onnekink C., Bloemers H.P.J., van de Ven W.J.M.; "Evolutionary conserved close linkage of the c-fes/fps proto-oncogene and genetic sequences encoding a receptor-like protein."; EMBO J. 5:2197-2202(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COFACTOR: CALCIUM-DEPENDENT.
ENZYME REGULATION: COULD BE INHIBITED BY THE NOT SECONDLY CLEAVED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDINE-9191619: PubMed-8020465;
Siezen R.J., Creemers J.W.M., van de Ven W.J.M.;
"Homology modelling of the catalytic domain of human furin. A model for the eukaryotic subtilish-like proprotein convertases.";
Eur. J. Blochem. 222:255-266(1994).
                                                                                                                                                                                                                                                                                                                                          "CDNA and gene structure for a human subtilisin-like protease with cleavage specificity for paired basic amino acid residues."; DNA Cell Biol. 10:319-328(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE SPECIFICITY: SEEMS TO BE EXPRESSED UBIQUITOUSLY.
DOMAIN: CONTAINS A CYTOPLANIC DOMAIN RESPONSTBLE FOR ITS TGN
LOCALIZATION AND RECYCLING FROM THE CELL SURRAGE.
FTM: THE PROPEPTIDE IS AUTOCATALYTICALLY REMOVED THROUGH AN
INTRAMOLECULAR CLEANAGE PROBABLY IN THE ENDOPLASMIC RETICULUM
(ER). IN THE TGN THE SECOND CLEAVAGE IN THE PROPEPTIDE COULD IT
TO THE ACTIVATION OF FURIN
                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY.
                                                                                                                                                                                                    "Structural homology between the human fur gene product and the subtilisin-like protease encoded by yeast KEX2."; Nucleic Acids Res. 18:664-664(1990).
                                                                                                                                                                                                                                                                                                               Barr P.J., Mason O.B., Landsberg K.E., Wong P.A., Kiefer M.C.,
                                                                                                                                        TISSUE-Blood;
MEDLINE-90175002; PubMed-2408021;
van den Ouweland A.M.W., van Duijnhoven H.L.P., Keizer G.D.,
Dorssers L.C.J., van de Ven W.J.M.;
cleaving enzyme) (PACE) (Dibasic processing enzyme). FURIN OR PACE OR FUR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3D-STRUCTURE MODELING OF CATALYTIC DOMAIN.
                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
MEDLINE=91321735; Pubmed=1713771;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   respective precursors.
                                   Homo sapiens (Human).
                                                                                                                        SEQUENCE FROM N.A.
                                                                                      NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                  Brake A.J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3888888888888888888888888888
```

```
the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SYSTEM (BY SIMILARITY).
SYSTEM (BY SIMILARITY).
SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hydrolase; Serine protease; Transmembrane; Glycoprotein; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC. . .) (POTENTIAL)
CLEAVAGE (SECOND AUTO-).
CLEAVAGE (FIRST AUTO-).
CELL ATTACHMENT SITE (POTENTIAL).
CELL SURFACE SIGNAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 794;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANS GOLGI NETWORK SIGNAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10C44DD5892EF35D CRC64;
                                                                                                                                                                                                                                             GO; GO:0005794; C:GOlgi apparatus; TAS.
GO; GO:0004276; F:furin activity; TAS.
GO; GO:0007267; P:cell-cell signaling; TAS.
GO; GO:0006508; P:proteclysis and peptidolysis; TAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ٠,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 55; DB 1;
Pred. No. 2.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHARGE RELAY S
CHARGE RELAY S
CHARGE RELAY S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mismatches
                                                                                                                                                                                                                                                                                                               InterPro: IPR006212; Furin_repeat.
InterPro: IPR002804; P. domain.
InterPro: IPR00209; Peptidase_S8.
Ffam; PF01403; P. Proprotein; PARTIAL.
Pfam; PF001082; Peptidase_S8; 1.
PRINTS; PR00723; SUBTILISIN.
ProDom; PD000717; P_domain; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROSITE; PS00136; SUBTILASE_ASP; PROSITE; PS00137; SUBTILASE_HIS; PROSITE; PS00138; SUBTILASE_SER;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         :|:|||: :: | | |||
102 KRRTKRDVYQEPTDPKFP 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 QRKTKRNAHRRPQDVKFP 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                86678 MW;
                                                                                                       EMBL; X17094; CAA34948.1; -. EMBL; X04329; CAA27860.1; -. EMBL; A06939; CAA00605.1; -. PIR; A39552; XXBUF. HSSP; Q99405; 1MPT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               35.7%;
50.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                          MEROPS; S08.071; -. Genew; HGNC:8568; FURIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                794 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Symogen; Calcium
                                                                                                                                                                                                                              MIM; 136950;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SITE
```

7, 2003, 11:20:01 Search completed: August Job time : 4.90909 secs

This SWISS-PROT entry is copyright. It is produced through a collaboration

hepatitis hepatitis

hepatitis nepatitis hepatitis hepatitis

ALIGNMENTS

hepatitis hepatitis

hepatit hepatit

Q68682 Q68708 Q68679 Q68683 Q68683 Q68707 Q68707 Q68785 Q87780 Q87781 Q81264

hepatitis hepatitis

hepatitis hepatitis

hepatitis hepatitis hepatitis hepatitis hepatitis

0068682 00688708 00688708 00688708 00688708 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781 0089781

080P52 080P88

```
000000000000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          P90364 hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          hepatitis
                                                          August 7, 2003, 11:05:41; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              099xn3
068575
068575
081980
081541
068311
068316
0819x9
0819x8
0819x7
0819x7
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                               stal number of hits satisfying chosen parameters:
                                                                                                   US-09-491-146A-28
154
1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                           830525 segs, 258052604 residues
                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                        - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QBJYR5
QBJYR7
QBJYR6
Q68712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q68575
Q91AU0
Q81989
Q81541
Q68311
Q68306
Q68309
Q8JYR9
                                                                                                                                                                                                                                                                                                                                                                                                 sp_plant:*
sp_rodent:*
sp_virus:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:*
                                                                                                                                             BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                    sp_human:*
sp_invertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                         sp_organelle:*
                                                                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                       sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                          sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB
                                                                                                                                                                                                                                                                                             SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                          sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3010
3010
3010
                                                                                                                                                                                                                                                                                                                                                                     sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Title:
Perfect score:
                                                                                                                                              Scoring table:
                                         OM protein
                                                                                                                          Sednence:
                                                                                                                                                                            Searched:
                                                                                                                                                                                                                                                                                             Database
                                                            Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Result
```

T 1  P40364 P90364; D1-MAY-1997 (TrEMBLEEL: 03, 01-NAY-1997 (TrEMBLEEL: 03, 01-O-MAY-1997 (TremBLAZE) (TremBLET (TremBLET (TremBLET (TremBLET (TremBLINE)) (TremBLINE) (TremBL		PRELIMINARY; PRT; 105 AA.	(TrEMBLrel. 03, Created)	rEMBLrel. 03, Last sequence update)	(TrEMBLrel. 22, Last annotation update)	(Genome polyprotein) (Fragment).	rus.	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;		.03;		N.A.		of nine genotypes of hepatitis C virus using PCR		Thesis (1995), Nagoya City University Medical School.		N.A.	MEDLINE=96305314; PubMed=8712927;	tami M., Saleh M.G., Orito E., Ohba KI., Wu RR.	Koide T., Tibbs C.J., Nouri-Aria K.T., Tokudome S., Williams R.;	"Usefulness and limitation of phylogenetic analysis for hepatitis C	virus core region: application to isolates from Egiptian and Yemeni		Arch. Virol. 141:1101-1113(1996).	SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY		PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF	PROTEIN C AND MRNA (BY SIMILARITY).	EMBL; D82034; BAA11519.1;	InterPro; IPR002522; HCV_capsid.	HCV_capsid; 1.		105 105
	RESULT 1 P90364		1997			Core protein (	Hepatitis C virus.	iruses; ssRNA	Hepacivirus.	CBI_TaxID=111	1]	SEQUENCE FROM N.A.	Ohno T., Mizokami M.;	"Determination of nine	method.";	hesis (1995),	2]	SEQUENCE FROM N.A.	EDLINE-963053	hno T., Mizok	oide T., Tibb	Usefulness an	irus core reg	atients.";	rch. Virol. 1	-!- SUBUNIT: T	LIPOPROTEI	PROTEIN M	PROTEIN C	MBL; D82034;	nterPro; IPR(	fam; PF01543;	Polyprotein.	

hepatitis hepatitis

us-09-491-146a-28.rspt

```
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE 191 AA;
                                             Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus
                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                         068575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              091AU0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        091AU0
                                                                                                                                                                                                                                                         a
                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus type 2. VIruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGNS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-94351179; PubMed-7520922;
Nakazawa T., Kato N., Ohkoshi S., Shibuya A., Shimotohno K.;
J. Hepatol. 20:623-629(1994).
J. Hepatol. 20:623-629(1994).
LIPOPROTEIN EVIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIXCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
INTERPORT IN PRO2522; HCV_capsid.
Promis PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ohno T., Mizokami M.; "Determination of nine genotypes of hepatitis C virus using PCR
                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 144; DB 12; Length 106;
Pred. No. 6.9e-14;
1; Mismatches 1; Indels C
Score 144; DB 12; Length 105;
Pred. No. 6.8e-14;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                          01-OCT-2002 (TrEMBLrel. 22, Last annotation update) (isolate USA8) genomic RNA (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     106
12001 MW; 25D0D5414B3EA9DC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           D945CB60661797E3 CRC64;
                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09PNN3;
01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update
01-CT-2002 (TrEMBLrel. 22, Last annotation update
01-CT-2002 (TrEMBLrel. 22, Last annotation update
EZ/NS1 protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                 106 AA
                                                                                  1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                          1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           125 AA; 14220 MW;
93.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ch 93.5%;
1 Similarity 92.9%;
26; Conservative
                                                                                                                                                                                                                                                                      01-NOV-1996 (TrEMBLrel. 01,
                   Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                           (TremBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       106 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-40271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein.
NON_TER 1
SEQUENCE 10
                                                                                                                                                                                                                                                                                           01-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-USA8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
SEQUENCE 1
  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              mehtod.
                                                                                                                                                                                                                               081831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9PXN3
                                                                                                                                                                                   RESULT 2
081831
1D 0818
AC 0818
DT 01-N
DT 01-O
DE (150
OS HEPA
OC VITU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SULT 3
                                                                                  ò
                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPRESENTATION OF THE PROPERTY OF THE PROPERTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Modecular cloning and expression of hepatitis Curus, core protein and production of monoclonal antibodies to the recombinant protein.", Asian Pac. J. Allergy Immunol. 14:0-0(1996).

-I SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROPEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF FROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; U55284; AABO015.1; -
INTERPO: IPRO02522; HCV_capsid.

InterPo: IPRO02521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Lim S.P., Khu Y.L., Hong W.J., Tay A., Ting A.E., Lim S.G., Tan Y.H.; "Identification and molecular characterization of the complete genome of a Singapore isolate of hepatitis C virus: sequence comparison with other strains and phylogenetic analysis."; virus Genes 23:89-95(2001)
                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 191;
     Length 125,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Songsivilai S., Dharakul T., Kunkitti R., Thepthai C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   191
20840 MW; 4AAE63444D8329E2 CRC64;
                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Last sequence update)
Last annotation update)
Score 144; DB 12;
Pred. No. 8.2e-14;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93.5%; Score 144; DB 12;
92.9%; Pred. No. 1.3e-13;
11ve 1; Mismatches 1;
                                                                                                                                                                                                                                                                                                    191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3010 AA.
                                                                                                                              5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                          1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Created)
                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=HCV-S1;
MEDLINE=21440119; Pubmed=11556407;
93.5%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF01543; HCV_capsid;
Pfam; PF01542; HCV_core; 1.
                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3est_Local Similarity 92.9
4atches 26; Conservative
```

```
InterPro; IPR001410; DEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEXDc;
                                                                                                                                                                                                                                                 IPR002531;
                                                                                                                                                                                                                                                                                                        IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                   IPR001650;
                                                                                                                                                                                                                                                                                                                          IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SM00487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF00998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF00271
                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                     nterPro:
                                                                                                                                                                                                                                                                                                                                                               interPro;
                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                          interPro;
                                                                                                                                                                                                                                                                                                                                                                                   interPro;
                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                        nterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  081541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ofam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              081541
 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SMART: SNOWLY, CALLOCK, PROSTILE, PS00190; CYTOCHROME_C; 1.
PROSTIE; PS50501; RDRP_POSITIVE; 1.
PROSTIE; PS50521; RDRP_VIRAL; 1.
Coat protein; Envelope protein; 1.
Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
SROUENCE 3010 AA; 326793 MW; 3D89304314F9F795 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Zheng W.Z.;
"Genotype identification of hepatitis c virus (HCV) isolated from a
                                                   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SUBMNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 12; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q81989;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAN-2003 (TrEMBLrel. 23, Last annotation update)
El and E2/NSI envelope glycoprotein (Genome polyprotein).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  93.5%; Score 144; DB 12; 92,9%; Pred. No. 2.4e-12; 1ive 1; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                 EMBL; AF356827; AAL00900.1; -.
InterPro; IPR000345; CytC_heme_bind.
                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_VIr.
Pfam: PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                      IPR002522; HCV_capsid.
IPR002521; HCV_core.
IPR002519; HCV_env.
                                                                                                                                                                                                                                                               IPR002531; HCV_NS1.
IPR004109; HCV_NS3.
IPR004109; HCV_NS3.
IPR000745; HCV_NS4a.
IPR001490; HCV_NS4b.
IPR002868; HCV_NS5a.
IPR002166; HCV_RGRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                        InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF02907; HCV_NS3; 1.
PF01006; HCV_NS4a; 1
PF01001; HCV_NS4b; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SMART; SM00487; DEXDC;
 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11103;
                 STRAIN-HCV-S1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-HCV-N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-HCV-N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Zheng W.-Y.Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Zheng W.Z.
                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                     Lim S.P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              081989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 6
Q81989
THE REPORT OF TH
```

```
single Japanese carrier in Nagasaki prefecture and genome analysis of El and E2/NS1 envelope glycoprotein regions.";
Jpn. J. Trop. Med. Hyg. 22:169-177(1994).
--- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL: D63857; BAA09919.1; --
HSSP; P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J., Nakazawa T., Hijikata M., Ishimura Y., Shimotohno K.; "Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among infected individuals.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        93.5%; Score 144; DB 12; Length 3010; 92.9%; Pred. No. 2.4e-12; ive 1; Mismatches 1; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3010 AA; 327503 MW; C7BDB38169D6E3CF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         081541;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPQRKTKRNTYRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR007095; RNA_pol_DS_PS. IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-92295714; PubMed-1318627;
                                                                                                                                                                                                                                                                                                                               HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS5a.
HCV_RdRP.
Helicase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P01001; HCV_NS4b; 1.
P010106; HCV_NS5a; 1.
P000271; helicase_C; 1.
P000998; Viral_RGRP; 1.
P0186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS1
```

Gaps

ö

Indels

Length 45;

Score 143; DB 12; Pred. No. 3.9e-14; 1; Mismatches 1;

CC527167096AAA81 CRC64;

```
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; U23749; AAA65056.1; -- InterPro; IPRO0252; HCV_capsid.
                                                                                                                                                                                                                                                                                                                   45
5015 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                     92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   45 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        45
45 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-HCV-BB5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein.
NON_TER .
SEQUENCE 45
                                                                                                                                                                                                                                                                                    Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hebacivirus
                                                                                                                                                                                                                                                                                                                   NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q68306
Q68306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     068309;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  068309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 9
Q68306
                    SHWENDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SO TENERAL SOCO SET THE SOCO SE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polymerase; Transferase; Transmembrane
VITUS RES. 23:39-53(1992).

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GIVCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; D1135; BAA1894.1; -.

HSSP; P26663; 1JXP.

InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-HCV-BB37;
Songsivilai S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROSITE: P550507; RDRP_POSITIVE; 1.
PROSITE; P550521; RDRP_VIRAL; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural;
POlyprotein; RNA directed RNA polymerase; Transferase; Transferase; Transferase; Transfer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            05F0B2102CF9DD9D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NOV-1996 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 144; DB 12;
Pred. No. 2.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          45 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. 2.46
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NS4A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR002166; HCV_RdRP.
IPR007095; RNA_pol_DS_PS.
IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                         IPRO0252; HCV_capsid.
IPRO02521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS5a
                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           93.5%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   F00998; Viral_RdRP; 1
PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF02907; HCV_NS3; 1.
PF01006; HCV_NS4a; 1.
PF01001; HCV_NS4b; 1.
PF01506; HCV_NS5a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            326564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01543; HCV_capsid;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_core;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3est_Local Similarity 92.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_env;
HCV_NS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1006
1615
1862
2013
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01538; HCV_NS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2014
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF00998
                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro:
                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              nterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         virus.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q68311
Q68311;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 8
068311
068311
068331
AC 06833
DD 01-N
DT 01-N
      CONTRACTOR OF THE PROPERTY OF 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Songsivilai S., Kanistanon D., Kunkitti R.; "Identification and characterisation of Thai isolates of hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN ENVELOPE.
PROFEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CC4C21ED236AAA81 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, Last Sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NVV-1996 (TrEMBLrel. 01, Last sequence update)
01-0CT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 143; DB 12;
Pred. No. 3.9e-14;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                     45 AA
1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                           5 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last seq
                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    U23744; AAA65051.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  45
5030 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 10
068309
AC 06830
DT 01-NO
DT 01-NO
DE COFE
OS VITUB
CO HEPAC
OC HE
```

```
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                     InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                         Genome polyprotein (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
6897 MW;
                                                                                                                                                                                              92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                          01-OCT-2002 (TrEMBLrel. 22, 01-OCT-2002 (TrEMBLrel. 22, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (TrEMBLrel. 22, (TrEMBLrel. 22,
                                                                                                                                                                                           92.9%
Best Local Similarity 92.9%
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
61 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
les 26; Conserv
                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                  NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11103;
                                                STRAIN-RIG286;
                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-RIG290;
                                                                                                                                                        Polyprotein.
NON_TER
SEQUENCE 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
NON_TER
SEQUENCE 61
                                                                                                                                                                                                                                                                                                                           01-OCT-2002
01-OCT-2002
                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-0CT-2002
01-0CT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                   QBJYR5
                                                                                                                                                                                                                                                                                                       Q8JYR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8JYR7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OBJYR7
                                                                                                                                                                                                                                                                                      RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
 <u>.</u>
                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEE
                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
        STRAIN-HCV-BB10;
Songsivilal S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis
                                            Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE.

LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A AND GINCORROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF EMBL; U23747; AAA55054.1;
--
InterPro; IPR00252; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 143; DB 12; Length 60;
Pred. No. 5.3e-14;
1; Mismatches 1; Indels
                                                                                                                                                                 DB 12; Length 46;
                                                                                                                                                                                     Indels
                                                                                                                                              ABDC4C21ED236AAA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60 AA; 6686 MW; B5691CBE8F26F2F1 CRC64;
                                                                                                                                                                                                                                                                                          01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (TrEMBLrel. 22, Created)
(TrEMBLrel. 22, Last sequence update)
(TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                    ;;
                                                                                                                                                                         Pred. No. 4e-14;
1; Mismatches
                                                                                                                                                                                                                                                                           60 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 AA
                                                                                                                                                                                                                 1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 143;
                                                                                                                                                                                                       1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                          Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Genome polyprotein (Fragment)
Hepatitis C virus.
                                                                                                                                                                92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           th 92.9%;
| Similarity 92.9%;
| Sonservative
                                                                                                                                              5129 MW;
                                                                                                                                                                Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-OCT-2002 (TrEMBLrel.
01-OCT-2002 (TrEMBLrel.
01-MAR-2003 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                              46 AA;
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-RIG306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
NON_TER
SEQUENCE 60
                                                                                                                          Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                       Hepacivirus
                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q8JYR8;
                                     virus.
                                                                                                                                                                                                                                                                          Q8JYR9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q8JYR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
Q8JYR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                      RESULT 11
ذ.
                                                                                                                                                                                                                         음
                                                                                                                                                                                                                                                                           STES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OE DI DI OS
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                               ď,
Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-I-SUBMINIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ВY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Shuston A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/Genbank/DDBJ databases.
-!-SIBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED ILIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX (PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; AF506620; AAM3396.1;
-InterPro; IPR002522; HCV_capsid.
PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                               Score 143; DB 12; Length 61;
Pred. No. 5.4e-14;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 143; DB 12; Length 61; Pred. No. 5.4e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                       61 61
61 AA; 6897 MW; EC656DC79E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EC656DC79E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Created)
Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
```

us-09-491-146a-28.rspt

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                           Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN E VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AF506618; AAM33394.1; -.
InterPro; IPR005522; HCV_capsid.
Pfam; PF01543; HCV_capsid. 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A
LIPOPROTEIN EVIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A
LIPOPROTEIN MAND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN AND GRAM BY SIMILARITY).

EMBL, AF506619; AAM33395.1; -
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid. 1.
                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                              Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                  Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 143; DB 12; Length 62;
Pred. No. 5.5e-14;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLYPFOCEIN.
NON_TER 62 62
SEQUENCE 62 AA; 7053 MW; EEEC656DC79E8F26 CRC64;
                                                                                                                                                                                                                                                                                                                                              EC657F706E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              v1-OCT-2002 (TrEMBLrel. 22, Created)
01-CCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Hepatitis C virus.
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                               Score 143; DB 12;
Pred. No. 5.4e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                             PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein.

NON_TER 61 61
SEQUENCE 61 AA; 6869 MW;
                                                                                                                                                                                                                                                                                                                                                                               92.98;
92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                            Ouery Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                 [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-RIG289;
                                                                                                                                STRAIN-RIG288;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 15
10 08JYR
AC 08JYR
AC 08JYR
DT 01-0C
DT 01-MA
DE Genom
OS Hepat
OC Virus
CO Nirus
RN ISI
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

Search completed: August 7, 2003, 11:19:01 Job time : 26.6364 secs

```
Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein HYPOTHETICAL: NO
 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ANTI-SENSE: NO
FRAGMENT TYPE: 1
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM:
                                                                                                                                                                                                                                                                                                               US-08-921-887-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-921-887-28
1, Appli
1, Appli
9, Appli
23, Appl
3, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  App
Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        , Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 28, Appl
Sequence 33, Appl
                                                                                      August 7, 2003, 11:07:41; Search time 10.5455 Seconds (without alignments) 112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 191, A Sequence 192, A Sequence 6, App
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 198,
Sequence 232,
Sequence 14, A
Sequence 8, Ap
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 2, Ap Sequence 1, Ap Sequence 1, Ap Sequence 1, Ap Sequence 2, Ap Sequence 3, Ap Sequence 26, Ap Sequence 23, Ap Sequence 3, Ap
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 10,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                         /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/1/iaa/pcTUS_COMB.pep:*
/cgn2_6/ptodata/1/iaa/pcTUS_COMB.pep:*
             GenCore version 5.1.6 . . Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-921-887-28
US-08-921-887-25
US-08-921-887-25
US-08-635-886C-191
US-08-88-160-6
US-08-38-160-1
US-09-389-756-1
US-09-389-756-1
US-09-389-756-1
US-08-380-160-1
US-08-380-160-1
US-08-380-160-1
US-08-380-160-1
US-08-380-160-1
US-08-380-160-1
US-08-380-160-1
US-08-380-160-1
US-08-380-160-1
US-08-330-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-836-075A-10
US-08-635-886C-198
US-08-635-886C-232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-836-075A-14
US-08-324-977-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-501-195-2
US-08-836-075A-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PCT-US93-08638-9
PCT-US94-07088-23
                                                                                                                                                                                                                                                            tal number of hits satisfying chosen parameters:
                                                                                                                                                        154
1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-384-616-8
                                                                                                                                                                                                                                  328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                protein search, using sw mode.
                                                                                                                                                                                            BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                              Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                     seq length: 0
seq length: 200000000
                                                                                                                                            US-09-491-146A-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ВВ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query
Match Length
                                                                                                                                                                                              Scoring table:
                                                                                                                                                        Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                                                                                                                                                                                                     Minimum DB
Maximum DB
                                                                                                                                                                                                                                                                                                                                                                                Database :
                                                                  protein
                                                                                                                                                                                                                                   Searched:
                                                                                           Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Result
```

```
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAN
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                             2, App. 16, Appl 1, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                             Sequence 2
Sequence 1
Sequence 1
                                                                                                                                                    Sequence 3
Sequence 3
Sequence 3
                                                                                                                                                                                                                 Sequence
Sequence
Sequence
Sequence
Sequence
Sequence
Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY: USA

ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: FIDPPY disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
FILING DATE:
CLASSIFICATION WIMBER: US/08/921,887
FILING DATE:
APPLICATION NUMBER: 03/08/921,887
FILING WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 36,714
REFERENCE/COCKET NUMBER: 36,714
RELEPRATION INFORMATION:
TELEFRAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LEMTH: 28 maning acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100.0%; Score 154; DB 3;
100.0%; Pred. No. 7.6e-16;
11ve 0; Mismatches 0;
                                                        PCT - USSS - 13552 - 2
US - 07 - 681 - 701 - 16
US - 08 - 078 - 271B - 11
US - 08 - 290 - 665A - 155
US - 08 - 290 - 665A - 155
US - 08 - 290 - 665A - 159
US - 08 - 290 - 665A - 160
US - 08 - 290 - 665A - 161
US - 08 - 290 - 665A - 161
US - 08 - 290 - 665A - 161
US - 08 - 290 - 665A - 164
US - 08 - 290 - 665A - 164
US - 08 - 290 - 665A - 164
US - 08 - 290 - 665A - 165
US - 08 - 290 - 665A - 165
US - 08 - 290 - 665A - 165
US - 08 - 290 - 665A - 165
                                                                                                                                                                                                                                                                                                                                                                                                                                       ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 28, Application US/08921887; Patent No. 6030771; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRANDEDNESS: not relevant
TOPOLOGY: not relevant
```

ò g

```
APPLICANT: LEROUX-ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VIRUS
FILE REFERENCE: 2752-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 450;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                              OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                    191 Peachtree Street, N.W., 37th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  93.5%; Score 144; DB 4; 92.9%; Pred. No. 4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           93.5%; Score 144; DB 3; illarity 92.9%; Pred. No. 2.1e-14; Conservative 1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CURRENT APPLICATION NUMBER: US/08/635,886C
CURRENT FILING DATE: 1996-04-25
PRIOR PRILORATION NUMBER: PCT/EP94/03555
PRIOR FILING DATE: 1994-10-28
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SEQ ID NO 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 191, Application US/08635886C
Patent No. 6555114
GENERAL INFORMATION:
                                                                             COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DG
                                                                                                                                                                                                                                                                                                                          ATTORNEY/AGENT INFORMATION:
NAME: WARKEN, WILLIAM L.
RECISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 0306
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEPAX: 404-818-3799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ) ORGANISM: Hepatitis virus US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: PRT
ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                   FILING DATE:
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches 26; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ANTI-SENSE: NO
FRAGMENT TYPE: 1
ORIGINAL SOURCE:
                                      Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-635-886C-191
ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                              APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 25, Application US/08921887

Patent No. 6030771

GENERAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 96.1%; Score 148; DE Best Local Similarity 96.4%; Pred. No. 5.6e Matches 27; Conservative 0; Mismatches
                                      PRPORKTKRNAHRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               03063-0380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                Sequence 33, Application US/08921887 Patent No. 6030771 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFRENCE/POCKET NUMBER: 0300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOLECULE TYPE: protein HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FRAGMENT TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ANTI-SENSE:
                                                                                                                                           US-08-921-887-33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-921-887-33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH:
```

g

0; Gaps

```
Length 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STATE: D.C.

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUW TYPE: Floppy disk

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PATENTIN Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/020,846

FILING DATE: 09-FEB-1998

CLASSIFICATION 1424

PRIOR APPLICATION NUMBER: 19 9-027015

FILING APPLICATION NUMBER: US/09-0200405

FILING APPLICATION NUMBER: US/09-0200405

FILING APPLICATION NUMBER: US/09-FEB-1996

ATOMNEY-MORENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 143; DB 3;
Pred. No. 3.7e-14;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION:
APPLICANT: YAMAGUCHI, Kenjiro
APPLICANT: RASHINAKUMA, Tomiko
APPLICANT: ATIBA, Yukie
APPLICANT: YAGI, Shintaro
APPLICANT: HASEGAMA, Akira
TITLE OF INVENTION: CHIMBRA ANTIGEN PEPTIDE
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 PKPQKKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: Wegner, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 053466/0225
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
WPB 28682
                                                                                                                                                                                                                                      TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORIGINAL SOURCE:
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 36, Application US/09020846 Patent No. 6322965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET: 3000 K Street, N.W. CITY: Washington
               TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                             92.9%;
92.9%;
REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                      LENGTH: 34 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                 INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                      TYPE: amino acid
STRANDEDNESS: Si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                             US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GENERAL INFORMATION:

APPLICANT: DALBON, PASCAL

APPLICANT: JOLIVET, Michel

TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE

TITLE OF INVENTION: HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY

TITLE OF INVENTION: FOR DETECTING THE LATTER

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF & BERRIDGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                 Sequence 192, Application US/08635886C
Patent No. 6555114
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT
APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: UNUS
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: IMMUSER: US/08/635,886C
CURRENT FILING DATE: 1996-04-25
PRIOR PLILING DATE: 1994-03555
PRIOR PLILING DATE: 1994-028
PRIOR PLILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTHAMRE: PATENTIN NOWER: DATE OF SEC ID NOS: 280
NUMBER OF SED ID NOS: 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 144; DB 4; Length 450; Pred, No. 4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 6, Application US/08380160
Patent No. 6235284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    17F: 2320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Floppy disk
COMPUTER: TIBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: Berridge, William P. REGISTRATION NUMBER: 30,024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 93.5%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT
ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE: OLIFF & BER
STREET: P.O. BOX 19928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CITY: Alexandria
STATE: VA
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
                                                                                                                                                                 US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 6
US-08-380-160-6
                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             á
```

```
Score 143; DB 3;
Pred. No. 4.8e-14;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 143; DB 4;
Pred. No. 4.8e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Mismatches
                                                                                1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                    4 PKPQRKTKRNTNRRPQDVKFPGGGOIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Patent No. 6576240
TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 88
                                                                                                                                                                                                                Sequence 1, Application US/09389756 Patent No. 6576240 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/08380160 Patent No. 6235284
      92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: Hepatitis C virus
Query Match
Best Local Similarity 92.9°
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: OLIFF & BERI
STREET: P.O. BOX 19928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RELEVANT RESIDUES: 2 TO 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PAGES: 3392-3396
DATE: 1991
                                                                                                                                                                                              JS-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: PRT
                                                                                à
                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                            RESULT 8
US-08-380-160-2
; Sequence 2, Application US/08380160
; Patent No. 6235284
; GENERAL INFORMATION:
APPLICANT: DALBON, Pascal
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: POR DETECTING THE LATTER
; TITLE OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                    ö
                                                                                              Length 43;
                                                                                                                                    1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER EADABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
COMPUTER: Patentin Release #1.0, Version #1.25
SOFTWARE: Patentin Release #1.0, Version #1.25
CURSERIATION NUMBER: US/08/380,160
FILING DATE:
CLASSIFFICATION: 530
PRIOR APPLICATION: DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 106-MAY: 1993
ATTORNEY/AGENT INFORMATION:
NAME: BETILIGE; W1111am P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WFB 28682
TELEPHONE: (70)3836-6400:
                                                                                            Score 143; DB 4;
Pred. No. 4.7e-14;
1; Mismatches 1;
                                                                                                                                                                                        1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Human Hepatitis C Virus STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BERRIDGE
                                                                                          Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: OLIFF & BERRI STREET: P.O. BOX 19928 CITY: Alexandria STATE: VA COUNTRY: USA ZIP: 22320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FRAGMENT TYPE: N-terminal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 44 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: (703)836-2787
              TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CTHER INFORMATION: /
CTHER INFORMATION: p
CTHER INFORMATION: p
CTHER INFORMATION: t.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY: Peptide LOCATION: 1..44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ANTI-SENSE: NO
STRANDEDNESS:
                                                                                                                                                                                                        QQ
                                                                                                                                                                      8
```

```
APPLICANT: JOLIVET, MICHEL
APPLICANT: DOLIVET, MICHEL
APPLICANT: DALBON, PRACOLS
APPLICANT: DALBON, PRACOLS
APPLICANT: DALBON, PASCAL
APPLICANT: LACONX, XAVIER
TILE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TILLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TILLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TILLE OF INVENTION: LACONY
TILLE OF INVENTION: LACONY
TILLE OF INVENTION: LACONY
TILLE OF INVENTION: PRACTICAL
TOWNER APPLICATION NUMBER: US/09/389,756
CURRENT FILING DATE: 1999-09-07
EARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION:

APPLICANT: DALBON, PASCA1

APPLICANT: DALBON, PASCA1

APPLICANT: DALBON, PASCA1

APPLICANT: DALBON, PASCA1

TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE

TITLE OF INVENTION: FOR DETECTING THE LATTER

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF & BERRIDGE
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PUBLICATION INFORMATION:
AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STATE: VA
COUNTRY: USA
ZIP: 22320
```

us-09-491-146a-28.rai

```
Sequence 23, Application US/08083947

Sequence 23, Application US/08083947

GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
APPLICANT: Hoseln, Barbara
TITLE OF INVENTION: in Diagnosing and Detecting No. 5639594-A, No. 5639594-
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                 Score 143; DB 1; Length 61;
Pred. No. 6.8e-14;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 143; DB 1; Length 61;
Pred. No. 6.8e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,947
FILING DATE: 19930628
CLASSIFICATION: 424
PRIOR APPLICATION DATE: US/08/054
FILING DATE: 15-SEP-1993
ATTORNEY/AGNIT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
  APPLICATION NUMBER: US/07/946,054 FILING DATE: 15-SEP-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34045
REFERENCE/DOCKET NUMBER: 20002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516)273-2828
                    FILING DATE: 15-SEP-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/POCKET NUMBER: 2000
TELECOMMUNICATION INFORMATION:
TELECHONE: 516-273-2828
TELEFRAX: 516-273-1177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92.98;
92.98;
                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO: 9: SEQUENCE CHARACTERISTICS: LENGTH: 61 amino acids TYPE: amino acid TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                     Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative 1
                                                                                                                                                                                                                                                                                                                  ; MOLECULE TYPE: peptide US-07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hauppauge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 , MOLECULE TYPE:
US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
APPLICANT: Work and Selection of Selection of Selection of TITLE OF INVENTION: No. 5582968-B Hepatitis
TITLE OF INVENTION: No. 5582968-B Hepatitis
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: United Blomedical Inc.
STREET: 25 Davids Dr.
CITY: Hauppauge
STATE: New York
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note- "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 143; DB 3; Length 45;
Pred. No. 4.9e-14;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 PKPORKTKRNTNRRPODVKFPGGGQIVG 32
                                                                                                                                                                                                                                                      NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
TELECOMMUNICATION:
TELEPHONE: (703)836-6400
TELEPRAX: (703)836-2787
                                                                                                                            FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ONGANISM: Human Hepatitis C Virus STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 9, Application US/07946054
Patent No. 5582968
                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FRAGMENT TYPE: N-terminal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CTHER INFORMATION:
COTHER INFORMATION:
COTHER INFORMATION:
CUS-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME/KEY: Peptide LOCATION: 1..45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ZIP: 11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ULT 11
```

ò

```
COMPUTER READABLE FORM: MEDIUM TYPE: FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: M. Lisa Wilson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 92.99
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: Amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PCT-US93-08638-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-26-37-26
Sequence 26, Application US/08262037
Sequence 26, Application US/08262037
Patent No. 5747239
GENERAL INFORMATION:
GENERAL INFORMATION:
TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
TITLE OF INVENTION: INFECTION AND PREVENTION THEREOF AS VACCINES
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVE.
         ö
       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                    GENERAL INFORMATION:
APPLICANT: Hoseln, Barbara
APPLICANT: Wang, Chang Yi
TILLE OF INVENTION: Peptides Effective for Diagnosis and
TITLE OF INVENTION: Detection of Hepatitis c Infection
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,550
     1; Mismatches
                                         1 PKPQRKTKRNAHRRPQDVKFPGGGGIVG 28
                                                          92.9%; Score 143; 92.9%; Pred. No. 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: W11son, M. L1sa
REGISTRATION NUMBER: 34,045
REFERENCE/POCKET NUMBER: 2000;
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516)273-2828
                                                                                                                                                                       Sequence 3, Application US/08530550 Patent No. 5736321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            uery Match 92.9
est Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      amino acid
                                                                                                                                                                                                                                                                                                                                                                             Hauppauge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                  USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                     US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                               STATE:
   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 14
                                         ò
                                                                          đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
CHANGER READER FORM
MEDIUM TYPE: LOOPY DISK
COMPATER INTO HEADER FORM
COMPATER INTO HEADER FORM
COMPATER READER FOR FOLDY
COMPATER READER FOLDY
COMPATER READER FOLDY
COMPATER READER FOLDY
CLASSIFICATION NUMBER: 07/19, 819
APPLICATION NUMBER: 07/66/275
APPLICATION NUMBER: 07/675
APPLICATION NUMBER: 07/6
```

ô

```
RECISTRATION NUMBER: 34,045

REFERENCE/DOCKET NUMBER: 9055

TELECOMMUNICATION INFORMATION:
TELEPANICATION INFORMATION:
TELEFAX: 516-273-2828

TELEFAX: 516-273-177

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acids
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08638-9
```

ö

0; Gaps

Search completed: August 7, 2003, 11:23:51 Job time : 10.6364 secs

Title: Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

Searched:

Database :

## ALIGNMENTS

APPLICANT: DOLIVET, MICHEL
APPLICANT: DOLIVET, MICHEL
APPLICANT: DOLIVET, MICHEL
APPLICANT: DALIVET, LANGOIS
APPLICANT: DALBON, PASCAL
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LACOUX, XAVIER
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: TREATING STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: TREATING AN HCV INFECTION
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILLE REPERENCE: 103999
CURRENT APPLICATION NUMBER: US/10/367,677
CURRENT FILING DATE: 1999-09-07
PRIOR FILING DATE: 1999-09-07
PRIOR FILING DATE: EARLIER APPLICATION NUMBER: PCT/FR98/00442
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PATENTIN VOY: 2.1
SSOFTWARE: PATENTIN VOY: 2.1 AUTHORS: Ogata, N. et al.

TITLE: Nuclectide Sequence and Mutation Rate of the H Strain
TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 88
PAGES: 3392-3396
DATE: 1991
RELEVANT RESIDUES: 2 TO 45 Sequence 1, Application US/10367677 Publication No. US20030118604A1 GENERAL INFORMATION: ORGANISM: Hepatitis C virus PUBLICATION INFORMATION: TYPE: PRT Query Match LENGTH Appl Appl Appl Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. Sequence 1, Sequence 1, Sequence 77, Sequence 14, Sequence 78, Sequence 66, Sequence 66, Sequence 65, Sequence 152, Sequence 152 Sequence 152, Sequence 42, Sequence 44, Sequence 42, Sequence 44, Sequence 1, Sequence 10, Description (GGGZ\_G) FORGATATA FUNDPARA/USOBA\_PUBGCOMB. PEP: \*
(GGGZ\_G/PUCCABLA/1/PUNDPARA/USOBA\_PUBGCOMB. PEP: \*
(GGGZ\_G/PUCCAGLA/1/PUNDPARA/USOBA\_PUBGCOMB. PEP: \*
(GGGZ\_G/PUCCAGLA/1/PUNDPARA/USOBA\_PUB. PEP: \*
(GGGZ\_G/PUCCAGLA/1/PUNDPARA/USIOA\_PUBGCOMB. PEP: \*
(GGGZ\_G/PUCCAGLA/1/PUNDPARA/USIOA\_PUBGCOMB. PEP: \*
(GGGZ\_G/PUCCAGLA/1/PUNDPARA/USIOA\_PUBGCOMB. PEP: \*
(GGGZ\_G/PUCCAGLA/1/PUNDPARA/USIOA\_PUB. PEP: \*
(GGGZ\_G/PUCCAGLA/1/PUNDPARA/USIOA\_PUB. PEP: \*
(GGGZ\_G/PUCCAGLA/1/PUNDPARA/USIOA\_PUB. PEP: \*
(GGGZ\_G/PUCCAGLA/1/PUNDPARA/USIOA\_PUB. PEP: \* /cgn2\_6/ptodata/1/pubpaa/USO7\_PUBCOMB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USO6\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USO6\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USO8\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USO8\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USO8\_NEW\_PUB.pep:\* US-09-756-875-8 US-09-921-397-77 US-09-821-138-14 US-09-851-138-46 US-09-851-138-46 US-09-871-138-46 US-09-878-281-152 US-09-878-281-152 US-09-899-046-42 US-09-899-046-44 US-09-878-281-44 US-10-367-677-1 US-09-851-138-10 US-09-758-308-1 SUMMARIES Match Length DB 137 138 166 169 169 169 117: 117: 117: 116: Result

Length 44;

Score 143; DB 15;

95.98;

Sequence 2, Appli Sequence 2, Appli Sequence 76, Appli Sequence 12, Appli Sequence 18, Appli Sequence 50, Appli Sequence 51, Appli Sequence 51, Appli Sequence 50, Appli	4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
US-09-929-955-2 US-10-104-966-2 US-09-194-496-3 US-09-194-496-3 US-09-851-138-16 US-09-851-138-18 US-09-891-148-18 US-09-891-146-55 US-09-891-046-54 US-09-878-281-52 US-09-878-281-52 US-09-878-281-52 US-09-978-281-54 US-09-991-55-50 US-09-991-55-50 US-09-991-305-50 US-09-991-305-50 US-09-991-305-50 US-09-991-305-50 US-09-991-305-50 US-09-991-305-50 US-09-991-305-50 US-09-991-305-50 US-09-991-305-50 US-09-991-305-50 US-09-991-305-50 US-09-991-305-50 US-09-991-305-50 US-09-991-305-50	US-10-259-275-40 US-09-742-659-4 US-09-922-552-5 US-09-929-955-1 US-09-747-419-20 US-09-891-894-3 US-10-104-966-1 US-10-259-275-20 US-09-995-937-2
044400000000000000000000000000000000000	
1188 3391 3311 3311 3311 3311 3311 3311	2985 3011 3011 3011 3011 3011 3011 3012
0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.	00000000000000000000000000000000000000
######################################	
	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 97;
                                                                                                                                                                                                                     Length 91;
                                                                                                                                                                                                                                                           1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ouery Match 92.9%; Score 143; DB 10; Best Local Similarity 92.9%; Pred. No. 1.3e-12; Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES NUMBER OF SEQUENCES: 29 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                   Score 143; DB 9;
Pred. No. 1.3e-12;
1; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-AUG-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATORNEY/AGENT INFORMATION:
NAME: ERNEY. BARBARA G.
REFERENCE/DOCKET NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1808-157A
CURRENT FILING DATE: 2001-01-10
PRICOR APPLICATION NUMBER: 60/092,
PRIOR FILING DATE: 1999-07-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PATENTIN VERSION 3.0
SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 8, Application US/09756875
Patent No. US20020150990A1
GENERAL INFORMATION:
APPLICANT: PIKE, IAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEPHONE: (202)783-6040 INFORMATION FOR SEQ ID NO: 8:
                                                                                                                                                                                                                 Ouery Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                      ; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY: unknown
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20004
                                                                                                                                                                                                                                                                                                                                                                                                           US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ŋ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EDITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 14114.0349U2
CURRENT APPLICATION NUMBER: US/09/758,308
                                                                                                                                                                                      Sequence 10, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION: GEERT
APPLICANT: MAERTENS, GEERT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AGENTS
AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                        Gaps
                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOFTWARE: Microsoft Word 6.0 / ASCII text output CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
Pred. No. 5.9e-13;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.9%; Score 143; DB 10; 92.9%; Pred. No. 1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: 08/836,075
FILING DATE: CUNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                   CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 10:
                                                        COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY: 11near
MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPORKTKRNAHRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                  STAILS. ... COUNTRY: USA ZIP: 77210-4433 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 1, Application US/09758308 Patent No. US20020090607A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS:
LENGTH: 74 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                           STREET: P.O. BOX 4433
92.98;
Best Local Similarity 92.99
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 92,9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                            CITY: HOUSTON
                                                                                                                                                                          US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 3
US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                        RESULT 2
                                                                                             q
                                                            ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    င်
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
```

```
Sequence 46, Application US/09851138
Sequence 46, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGN
                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                            polypeptides selected from
the hepatitis C virus and
                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                   Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 143; DB 10;
Pred, No. 1.6e-12;
                                                                   Score 143; DB 10;
Pred. No. 1.5e-12;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOFTWARE: Microsoft Word 6.0 / ASCII CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                             Sequence 78, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
TEDE OF INVENTION: SID nucleic acids and poly TITLE OF INVENTION: pathogenic strain of the FILE REFERENCE: B4809A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
PRIOR FILING DATE: 2000-08-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: US/09/851,138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                           1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                  1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; SEQUENCE DESCRIPTION: SEQ ID NO: 14: US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: 08/836,075
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE: 09-May-2001 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHIJ
SIREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AGENTS
                                                                   Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQ ID NOS: 156
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-851-138-46
                                                                                                                                                                                                                                                                                             US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
                                                                                                                                                           õ
                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HEPATITIS C VIRUS GENOTYPES
PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                        TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and FITLE OF INVENTION: applications thereof FILE REFERENCE: B4809A - JAZ CURRENT APPLICATION NUMBER: US/09/921,397 CURRENT FILING DATE: 2001-08-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
92.9%; .Score 143; DB 10;
Best Local Similarity 92.9%; Pred. No. 1.4e-12;
Matches 26; Conservative 1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 14. Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
STUTVER, LIEVENT
TITLE OF INVENTION: NEW SEQUENCES OF
AND THEIR USE AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                          CURRENT FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: EP 00402225.7
PRIOR FILING DATE: 2000-08-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                       Sequence 77, Application US/09921397 Patent No. US20020151484A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: Hepatitis C virus 9-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                           NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 77
LENGTH: 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CITY: HOUSTON STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: USA
                                                                     GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-851-138-14
  US-09-921-397-77
```

ò g

```
TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy. NUMBER OF SECURNES: 270 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TILLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COmpatible
                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 166;
                Indels
                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA:
              ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 143; DB 11;
Pred. No. 2.4e-12;
1; Mismatches 1
            Mismatches
                                                                      1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 PKPORKTKRNINRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92.9%; Score 143;
                                                 1 PKPORKTKRNAHRRPODVKFPGGGQIVG
                                                                                                                                                                                   Sequence 152, Application US/09899046; Publication No. US20030008274A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US/09/878,28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 152, Application US/09878281 Publication No. US20030032005A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO: 152:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: 152: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 92.99
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS:
          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      , MOLECULE TYPE: protein US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; MOLECULE TYPE:
US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-878-281-152
          26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
          Matches
                                                                                g
                                                 ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AFFLICANT STUYVER, LIEVEN
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: 08/836,075
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                       Length 137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 138;
                                                                                                                                                                                                                                                                                                                                                            1; Indels
                                                                                                                                                                                                                                                                                                                Score 143; DB 10;
Pred. No. 1.9e-12;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 143; DB 1
Pred, No. 2e-12;
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                     ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                  MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HOLECULE TYPE: peptide SEQUENCE DESCRIPTION: SEQ ID NO: 60: US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: KAMMERER, PATRICIA A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 60, Application US/09851138 publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                        LENGTH: 137 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 138 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET: P.O. BOX 4433
                                                                                                                                                      SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO: 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                 TYPE: amino acid
                                                                                                                                                                                                                                                                                                            Query Match 92.9°
Best Local Similarity 92.9°
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                         US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                        3
```

```
TITLE OF INVENTION: New sequences of heparitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIOM TYPE: RIPAPY disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PACEOTIN Release #1.0, Version #1.25 (EPO)
APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                             Gaps
                                           ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 169;
    Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 143; DB 11;
Pred. No. 2.4e-12;
1; Mismatches 1;
  Score 143; DB 11;
Pred. No. 2.4e-12;
                                           1; Mismatches
                                                                                    1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PKPORKTKRNTNRRPODVKFPGGGQIVG 32
                                                                                                                           5 PKPORKTKRNTNRRPODVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE OF INVENTION: New sequences of TITLE OF INVENTION: genotypes for di NUMBER OF SEQUENCES: 270 COMPUTER READABLE FORM: MEDIUM TYPE: FLOPPy disk COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                         US-09-878-281-42
Sequence 42, Application US/05878281
Publication No. US20030032005A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 44, Application US/09878281; Publication No. US20030032005A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : 169 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 169 amino acids TYPE: amino acid
    Query Match 92.99
Best Local Similarity 92.99
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: protein US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          , MOLECULE TYPE: protein US-09-878-281-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    inear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-878-281-44
                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                New sequences of hepatitis C virus
genotypes for diagnosis, prophylaxis and therapy.
270
                          Gaps
                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and the conformer of SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                       1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRET APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 143; DB 11;
Pred. No. 2.4e-12;
    Pred. No. 2.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Mismatches
                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S PKPORKTKRNTNRRPODVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                            US-09-899-046-42; Sequence 42, Application US/09899046; Publication No. US20030008274A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Sequence 44, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.9%;
92.9%;
92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : 169 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 169 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
      Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               , MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                        q
```

```
Query Match
92.9%; Score 143; DB 11; Length 169;
Best Local Similarity 92.9%; Pred. No. 2.4e-12;
Matches 26; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

QY Db Search completed: August 7, 2003, 12:01:13 Job time : 14.3636 secs

genome polyprotein hypothetical prote genome polyprotein genome polyprotein

polyprotein polyprotein

genome

polyprotein polyprotein - hepa

> genome genome

polypeptide genome polyp

polyprotein polyprotein polyprotein polyprotein polyprotein polyprotein

genome genome

genome

genome

```
genome polyprotein - hepatitis C virus (genotype 1, N7) (fragment)
N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 1, N7
A.Variety: genotype 1, N7
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C.Accession: 54134
R.van Doorn, L.J.; Kleter, G.E.M.; Brower, J.T.
submitted to the EMBL Data Library, January 1994
A.Reference number: S4134
A.Reference number: S4134
                                                                                                                                                                                                                                                                                                                                                                                                                                     N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N9
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41351
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41351
A;Accession: S41351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
A; Cross-references: EMBL: 229452; NID: 9443866; PIDN: CAA82590.1; PID: 9443867
A; Experimental source: genotype 1, N7
A; Experimental source: genotype 1, N7
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus (genotype 1, N9) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 139; DB 2; Length 11 Pred. No. 4.2e-11; 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A:Experimental source: genotype 1, N9
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein *status predicted AMT>
                                                                                                                                                                                                                                                                                                                                  ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 139;
                                                 PN0677
JQ0881
JQ0881
PC2219
GNWVTC
GNWVCJ
S18030
GNWVCH
                                                                                                                                                                                                                   JC5620
JQ1303
                                                                                                                                                                                                                                                          GNWVJ8
S41356
                                    518032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            89.1%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A) Molecule type: genomic RNA
A) Residues: 1-115 <VAN>
A) Cross-references: EMBL: 229454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                            3010
3010
3010
3011
3011
3033
108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 26; Conserv
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 2
S41349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 oolyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein
                                                                                                                        August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           депоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       <del>депо</del>ше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               депоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    депоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome
                                                                                                                                                                                                                                                                                                                                                                      283308
                 GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                   al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                   156
1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                  283308 segs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                         protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S41344
S41350
S41354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S41369
S41368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JQ1584
S19875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PC1284
JQ1925
                                                                                                                                                                                                                                                                                              Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S41342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    S41351
                                                                                                                                                                                                                                                                                                                                                                                                         length: 0
length: 2000000000
                                                                                                                                                                                                    US-09-491-146A-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ВВ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            115
118
369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
                                                                                                                                                                                                                                                                               BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         \vec{v} \vec{w} \vec{v} 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Post-processing:
                                                                                                                                                                                                                                                                                                                                                                                                           seq
                                                                                                                                                                                                                       Perfect score:
                                                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                         OM protein
                                                                                                                                                                                                                                                                                                                                                                                                           Minimum DB
Maximum DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database :
                                                                                                                                                                                                                                        Sequence:
                                                                                                                                                                                                                                                                                                                                  Searched:
                                                                                                                             Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š.
```

ö

Gaps

; 0

Length 115;

DB 2;

Length 115;

Best Local

Matches

g

õ

```
C. Accession: PC2060
R.Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
Biochem. Blophys. Res. Commun. 199, 1474-1481, 1994
A.Title: Identification of the third major genotype of hepatitis C virus in France
A.Reference number: PC2060, MUID:94197744; PMID:8147893
A.Recession: PC2060
A.Mocession: PC2060
A.Molecule type: mRNA
A.Molec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         submitted to the EMBL Data Library, April 1992
A; Description: A sentitive serodiagnosis of hepatitis C virus infection with two
A; Reference number: $21336
A; Accession: $21336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein S29 (core protein region) - hepatitis C virus (fragment) C;Species: hepatitis C virus (c;Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 17-Nov-2000 C;Accession: S21336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9cncme rorte protein
C.Species: hepatitis C virus
A:Variety: genotype 2, N2
C:Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                         N:Contains: envelope protein E1; nonstructural protein E2/NS1
C;Species: hepatitis C virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Molecule type: genomic RNA
A;Residues: 1-88 <SAT>
A;Cross-references: EMBL:X65548; NID:g59492; PIDN:CAA46517.1; PID:g59493
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus (genotype 2, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Accession: $41353
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 137; DB 2; Length 411;
Pred. No. 2.6e-10;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    87.2%; Score 136; DB 2; L6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A.Cross-references: EMBL:229456
A.Cross-references: EMBL:229456
A.Experimental source; genotype 2, N2
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG
                                                           polyprotein N1 - hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      87.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 89.33,
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: genomic RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Accession: S41353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 4
$19975
genome polyprotein - hepatitis C virus (isolate JK3) (fragment)
N;Contains: core protein; envelope protein 1; nonstructural protein 2; NSI/E2 protein
C;Species: hepatitis C virus
A;Variety: isolate JK3
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C;Accession: S19875
C;Accession: S19875
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
Scription: Sequence analysis of putative structural regions of Hepatitis C Virus isolatence number: S18029
A;Accession: S19875
A;Accession: S19875
A;Molecule type: genomic RNA
A;Residues: 1-782 <HON>
A;Accession: S19875
A;Molecule type: genomic RNA
A;Residues: 1-782 <HON>
A;Cross-references: EMBL:X61592; NID:959482; PIDN:CAA43789.1; PID:959483
A;Residues: core protein; envelope protein; envelope protein; onstructural
C;Superfamily: hepatitis C virus genome polyprotein
F;191/Product: core protein #status predicted <MAT2>
F;134-733/Product: envelope protein #status predicted <MAT3>
F;34-733/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>
                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
JQ1584
JQ01584
JQ01584
JQ01584
JQ01584
JQ01586
JQ01681
JQ0168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   predicted < (Asn) (cova
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  coss-references: GB:X84079; NID:g643119; PIDN:CAA58888.1; PID:g643120
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: core protein; envelope protein; glycoprotein; nonstructural protein; polypro
F;1-191/Product: core protein C *status predicted <CFC.>
F;192-389/Product: envelope protein E1 *status predicted <EE1>
F;390-640/Product: envelope protein E2 and nonstructural protein NSI *status predicted <F;196,209,234,305,417,430,448,476,540,556,576,623/Binding site: carbohydrate (Asn) (cove
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  putative
                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Title: Cloning and sequencing of the structural region and expression of A; Reference number: JQ1584; MUID:92300349; PMID:1318944
A; Accession: JQ1584
                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 640;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 782;
Pred. No. 4.2e-11;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 139; DB 2;
Pred. No. 2.2e-10;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88.5%; Score 138; DB 2;
92.9%; Pred. No. 3.6e-10;
11ve 0; Mismatches 2;
                                                                                                                                                                                           HILLI HILL HILLIHILLIHILI
PKPQRQTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HIFFIT HIFF THILL HITTELLING 32 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89.1%;
92.9%;
92.98;
Similarity 92.9
26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "olecule type: genomic RNA sidues: 1-640 < KUM>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
```

Matches

q

à

U

ö

Query Match Best Local Simi Matches 26;

qq

ö

genome

QQ

à

```
genome polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
N; Contains: core protein
C; Species: hepatitis C virus
C; Species: hepatitis C virus
A; Variety: genotype 1, N6
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 541348
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: 541341
A; Accession: 841348
A; Molecule type: genomic RNA
A; Residues: 1-108 < VANN
A; Residues: 1-108 < VANN
A; Cross-references: EMBL: Z29451
A; Cross-references: EMBL: Z29451
A; Experimental source: genotype 1, N6
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-108/product: core protein #status predicted < NATT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genome polyprotein - hepatitis C virus (genotype 1, NI) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, NI
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41341
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
S;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N5
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov
C;Accession: $41371
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
S;Ubmitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: $41371
A;Resion: $41371
A;Residues: 1-112 <VANN
A;Residues: 1-112 <VANN
A;Residues: 1-112 <VANN
A;Residues: 1-112 <VANN
C;Superimental source: genotype 5, N5
C;Superimental s
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 135; DB 2; Length 11
Pred. No. 1.3e-10;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 135; DB 2; L
Pred. No. 1.3e-10;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: genomic RNA A; Residues: 1-112 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Reference number: S41341
A; Accession: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A,Variety: genotype 2, N6
\ate: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           #text_change 17-Nov-2000
                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C.Species: hepatitis C. virus
A;Variety: genotype 2, N4
C;Species: hepatitis C. virus
A;Variety: genotype 2, N4
C;Date: 19-Way-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-200
C;Accession: S41355
an Doorn, L.J.: Kleter, G.E.M.; Brouwer, J.T.
nitted to the EMBL Data Library, January 1994
...Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41351
A;Accession: S41355
A;Molecule type: genomic RNA
A;Residues: 1-108 </br>
A;Cross-references: EMBL:229458
A;Experimental source: genotype 2, N4
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <AMT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cession: $41357

An Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.

Submitted to the EMBL Data Library, January 1994

A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.

A;Reference number: $41341
                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                               N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
                                                                         Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86.5%; Score 135; DB 2; I
89.3%; Pred. No. 1.3e-10;
ilve 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Experimental source: genotype 2, N6 C; Superfamily: hepatitis C vitus genome polyprotein C; Neywords: capsid protein; core protein; polyprotein F;1-108/Product: core protein *status predicted <MAT>
   F;1-108/Product: core protein #status predicted <MAT>
                                                                  Score 135; DB 2;
Pred. No. 1.3e-10;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 135; DB 2;
Pred. No. 1.3e-10;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            (genotype 2,
                                                                                                                                                                                                                                                     PKPORKTKRNTNRRPODVKFPGGGQIVG 32
                                                                                                                                                                                                                 1 PKPOROPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                               polyprotein - hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    86.5%;
89.3%;
                                                                     86.5%;
illarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
A; Cross-references: EMBL:229460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 89.39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 89.3°
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N;Contains: core protein
C;Species: hepatitis C virus
                                    Query Match
Best Local Similarity
Local 25; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N; Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Accession: S41357
```

ò g

LIPA.

ö

Gaps

ö

RESULT 10 S41348

á g

ö

Gaps

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               N; Contains: core protein
C; Species: hepatitis C virus
'lifety: genotype 5, N4
te: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
...cession: S41370
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
asubmitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LIPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein : hepatitis C virus (genotype 5, N3) (fragment)
».ontains: core protein
ecies: hepatitis C virus
ariety: genotype 5, N3
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A Molecule type: genomic RNA
A:Residues: 1-114 CVAN>
A:Residues: 1-114 CVAN>
A:Residues: 1-114 CVAN>
Cross-references: EMBL:229473; NID:g443908; PIDN:CAA82611.1; PID:g443909
A:Experimental source: genotype 5, N4
A:Experimental source: genotype 5, N4
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein
C:Reywords: capsid protein; core protein; polyprotein
F:1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: genomic RNA
A; Residues: 1-114 < VAN>
A; Residues: 1-114 < VAN>
A; Cross-references: EMBL:229472; NID:g443906; PIDN:CAA82610.1; PID:g443907
A; Experimental source: genotype 5, N3
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
A;Cross-references: EMBL:229444; NID:g443850; PIDN:CAA82582.1; PID:g443851
                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 114;
                                                                                                                                                           Length 112;
                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to
A; Reference number: S41341
                                                                                                                atus r-
Score 135; DB 2; Len
bred. No. 1.3e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 135; 'DB 2; L
Pred. No. 1.4e-10;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 86.5%; Score 135; DB 2; I
Best Local Similarity 89.3%; Pred. No. 1.4e-10;
Matches 25; Conservative 1; Mismatches 2;
                        A; Experimental source: genotypel, N1
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. 1.3e
1; Mismatches
                                                                                                                                                                                                                                                                                            1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPORQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 86.5%;
89.3%;
                                                                                                                                                        Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity 89.3.
The 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Accession: S41369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: S41369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

a

Search completed: August 7, 2003, 11:21:48 Job time : 10.5455 secs

```
ö
                  N; Contains: Core protein
C; Species: hepatitis C virus
A; Variety: genotype 5, N2
C; Date: 19-May.1994 # sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41368
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                                                                                                                                                                                                                                 A;Cross-references: EMBL:229471; NID:9443904; PIDN:CAA82609.1; PID:9443905 - A;Experimental source: genotype 5, N2 C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: capaid protein; core protein; polyprotein F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
C virus (genotype 5, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Match 86.5%; Score 135; DB 2; 1 Local Similarity 89.3%; Pred. No. 1.4e-10; les 25; Conservatiye 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
- hepatitis
                                                                                                                                                                                                                                                                                   A; Molecule type: genomic RNA A; Residues: 1-114 <VAN>
                                                                                                                                                                                                                                     A; Reference number: S41341
                                                                                                                                                                                                                                                             A; Accession: S41368
  genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local S
Matches 25,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dp
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec Run on:

US-09-491-146A-27 156 I PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched:

127863 al number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Description			Q01403 hepatitis c		hepatitis	Þ	h genome	모	c	æ	æ	c	Þ	2	2									_		Q49378 mycoplasma	1727	3185 sacchar	277	384 heliant	8800 dictyos	94933 homo sap	Q14738 h serine/th
ΩI	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVJT	POLG_HCV1	POLG_HCVTW	FMR2_HUMAN	GDF6_BOVIN	2A5D_RABIT	CSP_PLABE	Ω,	HOR7_HORVU	YLL2_EBV	YLL2_EBVA8	ABA2_HUMAN	MRP_STRSU	FMR2_MOUSE	RNB_HSV11	P32_MYCGA	GDA7_WHEAT	- 1	- 1	- 1	- 1	8	2A5D_HUMAN
DB	ч	-	-	7	~																				7	1	-	-		-	Н	Η.	Н
% Query Match Length	_	2	520	'n	737	3010	3010	3011	3033	3033	3010	3011	3010	1311	436	286	339	347	105	148	148	396	1256	1272	161	$\overline{}$	31	1064	338	493	511	977	602
& Ouery Match	9	86.5	86.5	ů.	ė.	ė.	86.5	ė.	ė.	è.	'n.	4.	'n	ä	σ	9	4.	4.	ش	ω.	ش	۳.	۳.	۳.	3	m.	Э.	m.	ς.	ά.	ä	32.7	32.4
Score	m	3	135	3	3	$\sim$	r	Š	$\sim$	3	'n	131	2	65	9	57.5	54	S	52.5	52	52	52	52	S	51.5				51	51	21	ω,	50.5
Result No.	н	7	е	4	2	9	7	80	0	10	11	12	13	14	15	16	17	18	19	50	21	22	23	24	25	26	27	28	59	30	31	32	33

P48608 drosophila P75136 mycoplasma	Q9qyx7 mus musculu P10164 rattus norv	P50477 canavalta e P10562 canavalta q	P14196 dictyosteli 035144 mus musculu	P10723 brugia mala Q92777 homo sapien	015297 homo sapien P21188 xenopus lae	
DIA_DROME YG55_MYCPN	PCLO_MOUSE PRP2_RAT	CANA_CANEN CANA_CANGL	AAC2_DICDI TRF2_MOUSE	SYN_BRUMA SYN2_HUMAN	P2CD_HUMAN ATND_XENLA	
1091	5038 172	445 445	448 495	548 582	605 277	
32.4	31.7	31.4	31.4 31.4	31.4 31.4	31.4 31.1	
50.5	49.5 49	4 4 0 0	4 4 0 0	<b>4.4</b> 0.0	49	
34 35	36 37	8 6 8 6	4 4 1	4 4 3 8	<b>44</b> 45	

## ALIGNMENTS

Gaps

ö

Indels

Length 520;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
BY SIMILARITY.
N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annoctation update)
16-OCT-2001 (Rel. 40, Last annoctation update)
Genome polyprotein (Contains: Capsid protein C (Core protein (P22);
Envelope glycoprotein El (GP32) (GP35), Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment).
Hepartis C virus (isolate HCV-KF) (HCV).
Hepartis C virus (isolate HCV-KF) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLING-93019030; pubMed=1383400;
Abe K., Inchauspe G., Fujisawa K.;
"Genomic characterization and mutation rate of hepatitis C virus
"Genomic characterization and mutation rate of hepatitis of nor-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992)
I. Gen. Virol. 73:2725-2729(1992)
I. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
ILIPOPPOTEIN BAYELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR: J01925; J01925.

InterPro: IPR002522; HCV_capsid.

InterPro: IPR002521; HCV_core.

InterPro: IPR002519; HCV_core.

InterPro: IPR002519; HCV_env.

InterPro: IPR002519; HCV_env.

Iffam: PF01543; HCV_capsid; 1.

Pfam: PF01540; HCV_core; 1.

Pfam: PF01550; HCV_env; 1.

Probom: PF01560; HCV_NS1; 1.

Probom: PF01660; HCV_NS1; 1.

Probom: Pf0186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                      56499 MW; AA135246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                  Score 135; DB 1; Lu
Pred. No. 1.4e-10;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         520 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Transmembrane; Nonstructural protein
                                                                                                                                                                                                                                                                                                                                                                                                  86.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; D10687; BAA01529.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 89.3
Les 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                      520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=31644;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVHK
Q01403;
                                                                                                                                                                                                                                                                                                                               NON_TER
SEQUENCE
                                                                                                      TRANSMEM
                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                         CARBOHYD
                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
POLG_HCVHK
                                                                                    CHAIN
          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-GOT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1) (Fragment).
Hepatitis C virus (isolate HCV-476) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flavivirian.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              K
                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-93019030; PubMed-1383400;
ADE K., Inchauspe G., Fujisawa K.;
ADE K., Inchauspe G., Fujisawa K.;
Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPMOTEIN BYMELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF01539; nc.________Pfam; PF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Probom; D0186062; HCV_LNS1; 1.
Probyprotein; G1ycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
Transmembrane; Nonstructural protein.
CELLULAR AMINOPEPTIDASE.
       NONSTRUCTURAL PROTEIN NS1 (POTENTIAL)
                                                                                                                        (POTENTIAL)
(POTENTIAL)
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                        86.5%; Score 135; DB 1; Length 513; 89.3%; Pred. No. 1.3e-10; Live 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                          943F31E3514CDEF3 CRC64;
                                                N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                520 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002522, HCV capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                          55704 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D10688; BAA01530.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                          513 AA;
                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                           25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVH4
                                                     CARBOHYD
CARBOHYD
                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                 NON_TER
SEQUENCE
                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                           Query Match
                              TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVH4
ID POLG_HCVH4
DT 01-JUL.
DT 01-JUL.
DE GENOME
DE ENVELOI
OC VITUSEI
OC VITUSEI
OC NEPACII
OC NE
                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FFF WWW DRAW DRAW DRAW WAR WAS DRAW DRAW WAS DRAW DRAW WAS DRAW WA
```

K

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         way
                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                     MATRIX PROPEIN C (FOLENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
BY SIMILARITY.
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural proteins NSI and NS2] (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         There are no restrictions on g as its content is in no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        K
                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus (isolate HC-J5) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     having poor
four distinct
                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                            Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-92230232; PubMed-1314459; Okamamoto K., Lizuka H., Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; Fill-length sequence of a hepatitis C virus genome having homology to reported isolates: comparative study of four genotypes.";
                                                                                                                                                                                                                                                       Score 135; DB 1; Length Dred, No. 1.4e-10; Indels
                                                                                                                                                                                                                                               1D2BD0A6FF27349B CRC64;
                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                       737 AA
                                                                                                                                                                                                                                                                                                       1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       European Bioinformatics Institute. The by non-profit institutions as long
                                                                                                                                                                                                                                                                                                                                   1 PKPOROPKRNTPRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Interpro; IPR002522; HCV_capsid.
Interpro; IPR002513; HCV_core.
Interpro; IPR002519; HCV_env.
Interpro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                .
Σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D10075; BAA00969.1; -.
                                                                                                                                                                                                                                                                          86.5%;
89.3%;
                                                                                                                                                                                                                                                56476
                                                                                                                                                                                                                                                                            Nuery Match 86.5 sest Local Similarity 89.3 watches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                233
234
305
418
                                                                                                                                                                                                                                               520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                       POLG_HCVJ5
                                                                                                                                                                                                                                 NON_TER
SEQUENCE
                                                                                                                                 CARBOHYD
                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                  CARBOHYD
 INIT_MET
                                                                                       TRANSMEM
                                                                                                    CARBOHYD
                                                                                                                   CARBOHYD
                                                                                                                                             CARBOHYD
                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                              윱
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                 CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ/EZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSZ (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
Hepatitis C virus (1solate HC-J7) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VITOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NSZA, NSZB, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-92230232; PubMed-1314459; Okamamoto K., Lizuka H., Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes.";
                                                                                                                                                                      REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 135; DB 1; Length 737;
Pred. No. 2e-10;
                                                                                                                           protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2; Indels
                                                                                                                                                                                                  CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3AF699D82AD501B1
                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. 2e-]
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
                                                                                                                                               Nonstructural protein.
Pfam; PF01539; HCV_core; 1.

Pfam; PF01560; HCV_NS1; 1.

Probom; PD186062; HCV_NS1; 1.

Polyprotein; Glycoprotein; Cat pr.
Transmembrane; Nonstructural and INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .
E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 89.3%; es 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          86.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      737
81207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCVJ7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                          TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     P27961
                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DDTTDDDDTTDDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDT
                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     요
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.lsb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VIruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                            REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    86.5%; Score 135; DB 1; Length 737; 89.3%; Pred. No. 2e-10; ive 1; Mismatches 2; Indels
                                                                                                                                                                                                                                      CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
                                                                                                                                                                                                       Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                          LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPOROPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                              POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LINKED
                                                                                                                                                                                                                  Nonstructural protein.
                                                                             EMBL, D10077; BAA00971.1; -.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PP01543; HCV_capsid; 1.
Pfam; PP01543; HCV_core; 1.
Pfam; PP01559; HCV_core; 1.
                                                                                                                                                                   Probom; PD18605 HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Col Transmembrane; Nonstructure:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 89.3
Matches 25, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      649
737
737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVBK
                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S
                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                            TRANSMEM
                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 6
POLG_HCVBK
                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-9827846; PubMed-956889;

MEDLINE-9827846; PubMed-956889;

Yan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,

Asteinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;

"Complex of NS3 protease and NS4A peptide of BK strain hepatitis C

virus; a 2.2-A resolution structure in a hexagonal crystal form.";

protein Sci. 7:837-847(1998).

"I Protein Sci. 7:83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 [RNA](N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                       Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; "Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   đ
                                                                                                                                                                                                                                                                                                                                                                                                               X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.

BODELINE-9701508B; Pubmed-8661916;

LOVE R.A., Parge H. E., Wickersham J.A., Hostomsky Z., Habuka N.,

MOOMRAW E.W., Adachi T., Hostomska Z.;

The crystral structure of hepatitis C virus NS3 proteinase reveals

trypsin-like fold and a structural zinc binding site.";
                                                                                                                                                                                                                            SEQUENCE OF 1487-1500.
MEDLINE-96235224; PubMed-8647104;
MEDLINE-96235224; PubMed-8647104;
BOTOWST P., Helland M., Oehlmann K., Becker B., Korneteky L.;
Non-structural protein 3 of hepetitis C virus inhibits
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Blochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                              MEDLINE-91140698; PubMed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_core.
                                                                                                                                                       from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A38465; GNWVTC.
1A1Q; 25-MAR-98.
1JXP; 14-JAN-98.
1NS3; 08-APR-98.
1C2P; 15-NOV-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26-JUN-00.
20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            interPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        nterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-APR-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS; S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1GX6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEROPS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PDB;
PDB;
PDB;
PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB;
```

us-09-491-146a-27.rsp

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 34, Last annotation update)
Genome polyprotein (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein E2 (GP63) (GP32); Envelope glycoprotein E2 (GP63) (GP32); Envelope glycoprotein E2 (GP63) (GP32); Envelope glycoprotein E2 (GP63) (GP63); Nostructural protein NSA (EC 3 4.22.); Protease/hallcase NS3 (P70) (Hepbolylrin)
(EC 3 4.22.98); Nonstructural protein NSA (P65); Nonstructural protein NSA (P65); (P70) (RNA-directed RNA POlymerase) (EC 2.77.48)]
Hepatitis C virus (isolate Japanese) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus (isolate Japanese) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  {RNA}(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate
                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                         Score 135; DB 1; Length 3010;
Pred. No. 8.9e-10;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-91088550; PubMed-2175903;
MEDLINE-91088550; PubMed-2175903;
Kato, Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi
Sugimura T., Shimotohno K.;
"Molecular Cloning of the human hepatitis C virus genome
Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                            327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPOROPKRNTPRRPODVKFPGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                  86.5%;
                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 89.33
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
1112
11120
11120
11133
11144
11161
11163
11171
11198
11198
11198
                                                                                                                                                                                                                                                                                        1680 168
3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11116;
                                                                                                                                                                                                    1175
1187
1189
1198
                                  1122
1129
1139
1139
11162
1162
1168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVJA
                                                                                                                                                                                                                                                                                          STRAND
SEQUENCE
                                                                                                                                                    TURN
STRAND
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
TURN
                                                                                STRAND
STRAND
HELIX
TURN
                                                                                                                                                                                                                                          STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P26662
g
                                                                                                                                                                                                                                                                                                                                                                       REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E, POTENTIAL).
NONSTRUCTURAL PROTEIN NS1, PEZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2, POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                                                            Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
          InterPro; IFR001499; HCV_NS4b.
InterPro; IFR001499; HCV_NS4b.
InterPro; IFR001868; HCV_NS5a.
InterPro; IFR0010166; HCV_NS5b.
InterPro; IFR001095; RNA_POL_DS_PS.
InterPro; IFR007095; RNA_POL_DS_PS.
InterPro; IFR007095; RNA_POL_DS_PS.
InterPro; IFR007094; RNA_POL_DS_PS.
InterPro; IFR007094; RNA_POL_DS_PS.
InterPro; IFR007094; RNA_POL_DS_PS.
InterPro; IFR007094; HCV_NS2; I.
Pfam; PF01907; HCV_NS3; I.
Pfam; PF01001; HCV_NS4b; I.
Pfam; PF01001; HCV_NS4b; I.
Pfam; PF00999; VITAL_RORP; I.
ProDom; P10186062; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          035
047
050
050
063
074
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            081
085
087
                                                                                                                                                                                                                                                                                                                                                              3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HELIX
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
```

Gaps

```
01-AUG-1992 (Rel. 23, Lest aequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Lest annotation update)
16-SEP-2003 (Rel. 42, Lest annotation Electric Elect
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-97331322; PubMed-9187654; Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.; Structure of the hepatitis C virus RNA helicase domain."; Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A., Murcko M.A., Lin C., Caron P.R.; "Hepatitis C virus NS3 RNA helicase domain with a bound oligonuclectide: the crystal structure provides insights into the mode
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.

CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACTIVATION OF NS3.
-I- FUNCTION: NS5A SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-I- FUNCTION: NS5B IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN ESSENTIAL ROLE IN THE VIRUS REPLICATION.
                                 (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
            (POTENTIAL)
                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                  Score 135; DB 1; Length 3010;
Pred. No. 8.9e-10;
l; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff
                                                                                                                                                                                                       MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-98154321; Pubmed-9493270;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657
                                                                                                      (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                          (GLCNAC.
                                                                                    (GLCNAC.
                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                          5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                 1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
       N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. MEDLINE-92052256; Pubmed-1658800;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-AUG-1992 (Rel. 23, Created)
                                           623
45 645
1 2041
2077
2240
2788
\; 327017 M
                                                                                                                                                                                                                                                     86.5%;
                                                                                                                                                                                                                                                                            89.3%;
                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                       Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus.
NCBI_TaxID=11108;
                                                                            645
2041
2077
2240
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    [RNA] (N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVH
P27958;
     CARBOHYD
CARBOHYD
CARBOHYD
                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                               CARBOHYD
                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 8
       821111118
                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DDT TO DDE BENDE B
                                                                                                                                                                                                                                                                                                                                                   à
                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CAPSID PROTEIN C (POTENTIAL).
MAJREY PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN NS! (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEBAEZ/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                      SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR00166; HCV_RdRP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
>fam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                           HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
HCV_NS2.
HCV_NS4.
HCV_NS4a.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam: PF02907; HCV_NS3; 1.
Pfam: PF01006; HCV_NS48; 1.
Pfam: PF01001; HCV_NS48; 1.
Pfam: PF01506; HCV_NS58; 1.
Pfam: PF00291; hellcase_C; 1.
Pfam: PF00998; VITAL_RGRP: 1.
                                                                                                                                                                                                                                                            EMBL; D90208; BAA14233.1;
PIR; A39253; GNWVCJ.
HSSP; P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF01543; HCV_capsid
PF01542; HCV_core;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SM00487; DEXDC;
PROTEIN C AND MRNA
                                                                                                                                                                                                                                                                                                                                      MEROPS; S29.001; -. MEROPS; U39.001; -. Interpro; IPR001410; D InterPro; IPR002522; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001490;
InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002518;
InterPro; IPR004109;
InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1316
196
209
234
250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF01539; PF01560; PF01538; P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PF02907; PF01006; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 347
1083
1107
1165
1230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 305
417
423
430
448
532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
```

DAR WANTER TO SERVICE TO SERVICE

us-09-491-146a-27.rsp

\$\$\$\$\$\$\$\$\$\$\$\$\$\$\$\$\$\$

N-LINKED (GLCNAC) (POTENTIAL)	2 MW; 772CBB29CCD94753 CRC64;
100 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	⋖
FT CARBOHYD 209 FT CARBOHYD 201 FT CARBOHYD 201 FT CARBOHYD 201 FT CARBOHYD 412 FT CARBOHYD 412 FT CARBOHYD 416 FT CARBOHYD 416 FT CARBOHYD 512 FT CARBOHYD 512 FT CARBOHYD 526 FT TURN 1221 FT TURN 1221 FT TURN 1221 FT TURN 1232 FT TURN 1310 FT TUR	SEQUENCE
1- SUBLIMIT: THE VIRION OF THIS VIRIS IS A NUCLECOAPSID COVERED BY A ADD E2. THOPPROTEINS ELL AND REAL.  AND E2. THOPPROTEINS THE NUCLECOARSID IS A COMPLEX OF PROTEINS. ELL AND REAL.  1- SHALLARITY: THE NGS PROTEINS NOT NOT BE ALL AND MENA.  1- SHALLARITY: THE NGS PROTEINS BELONGS TO PERTIDIASE PARILY STORMAND.  1- SHALLARITY: THE NGS PROTEINS BELONGS TO PERTIDIASE PARILY STORMAND.  1- SHALLARITY: THE NGS PROTEINS BELONGS TO PERTIDIASE PARILY STORMAND.  1- SHALLARITY: THE NGS PROTEINS BELONGS TO PERTIDIASE PARILY STORMAND.  1- SHALLARITY: THE NGS PROTEINS BELONGS TO PERTIDIASE PARILY STORMAND.  1- SHALLARITY: THE NGS PROTEINS BELONGS TO PERTIDIASE PARILY STORMAND.  1- SHALLARITY: THE NGS PROTEINS BELONGS TO PERTIDIASE PARILY STORMAND.  1- SHALLARITY: THE NGS PROTEINS BELONGS TO PERTIDIASE PARILY STORMAND.  1- SHALLARITY: THE NGS PROTEINS BELONGS TO PERTIDIASE PARILY STORMAND.  1- SHALLARITY: THE NGS PROTEINS BELONGS TO PERTIDIASE PARILY STORMAND.  1- SHALLARITY: THE NGS PROTEINS BELONGS TO PERTIDIASE PARILY STORMAND.  1- SHALLARITY: THE NGS PROTEINS AS I CAN BE THE OFFICE AND	1316 1319 DECH

us-09-491-146a-27.rsp

```
TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
g
                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=9204440; PubMed=1658196;

MEDLINE=9204440; PubMed=1658196;

MEDLINE=9204440; PubMed=1658196;

MEDLINE=9204440; PubMed=1658196;

MEDLINE=9204440; PubMed=1658196;

MACHODA A., Miyakawa Y., Mayumi M.;

Mucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions ";

J. Gen. Virol. 72:2697-2704(1991).

J. Gen. Virol. 72:2697-2704(1991).

HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

C. TATALITIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in pl and Ser or Ala in pl'.

C. CATALITIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                               1-2000.7

10.40G-1992 (Rel. 23, Created)

10.40G-1992 (Rel. 23, Last sequence update)

10.40G-1992 (Rel. 23, Last sequence update)

10.40G-1992 (Rel. 21, Last annoctation update)

10.40G-1992 (Rel. 41, Last annoctation (Relation of Legent (Rel. 41, Rel. 42, Rel.)

10.40G-1992 (Rel. 41, Rel. 41, R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (RNA)(N).

SUBURIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS.
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                        Gaps
                                                      ö
      Score 135; DB 1; Length 3011;
Pred. No. 8.9e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                  2; Indels
                                                                                                                                                                                                                                     PRT; 3033 AA
                                                    1; Mismatches
          86.5%; Score 135;
89.3%; Pred. No. 8
                                                                                        1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                               HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D00944; BAA00792.1;
                           Local Similarity 89.3
hes 25; Conservative
                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PIR; JQ1303; JQ1303.
HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002519;
InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000745;
InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11113;
                                                                                                                                                                                                                             POLG_HCVJ6
P26660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
          Query Match
                                                  Matches
                                                                                                                            a
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Core protein, Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE. CHAIN 116 191 MATRIX PROTEIN (POTENTIAL). CHAIN 192 383 MAJOR ENVELOPE PROTEIN E (POTENTIAL). CHAIN 384 733 NONSTRUCTURAL PROTEIN NSI (POTENTIAL). CHAIN 734 1010 NONSTRUCTURAL PROTEIN NSI (POTENTIAL). CHAIN 1011 1619 PROTEASE/HELICASE NSI (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       329165 MW; F957F5C1A273BE9E CRC64:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86.5%; Score 135; DB 1; 89.3%; Pred. No. 9e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (Rel. 23, Created)
(Rel. 23, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
| PPO; | IPNO. | IPNO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (N-LINKED (N-LINKED)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity 89.3
nes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             649
1091
2038
2811
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 10
POLG_HCVJ8
ID POLG_HCVJ8
AC P26661;
DT 01-AUG-1992 (
DT 01-AUG-1992 (
```

Fri Aug

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce.or send an email to license@isp-sib.ch).
Genome polyprotein (Contains: Capsid protein (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein El (GP31); Protein P7; Nonstructural protein NS2 (P20) (Hepacivirin) (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS48 (P27); Nonstructural protein NS58 (P56); Nonstructural protein NS58 (P66); P00) (RNA-directed RNA polymerase) (EC 2.7.7.48)]. Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                       MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                        the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A SUBUNIT: THE UNITION OF THIS VIRUS IS A NUCLEOCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                            VICOLOGY 188:331-341(1992).

-1- FUNCTION: THE SWALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDODHOBIC. SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viprecursor polyprotein. commonly with Asp or Glu in the P6 position, Cys or Thi in Pl and Ser or Ala in Pl'.

-1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate
                                                                                                                                                                                                                                                                                                         *Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -1 - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002166; HCV_RGRP.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_caps1d.
HCV_core.
HCV_core.
HCV_NS1.
HCV_NS3.
HCV_NS3.
HCV_NS4.
HCV_NS4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D10988; BAA01761.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01506; HCV_NS5a; 1.
PF00998; V1ral_RdRP; 1
n; PD186062; HCV_NS1; 1
SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01538; HCV_NS2; 1.

; PF02907; HCV_NS3; 1.

; PF01006; HCV_NS4a; 1.

; PF01001; HCV_NS4b; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01539; HCV_env;
PF01560; HCV_NS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A40250; GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR002521:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            nterPro; IPR004109
                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                    NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .00'6En
                                                                                                                                                                                                                                                                                                                                                 genotypes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ProDom;
SMART; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               fam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   fam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                              ,555555555555555555555555555
```

```
CELLULAR ANINOPERTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1; Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MW; 1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-92295714; PubMed-1318627;
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
                                                                                                                                                                                                                                              SYSTEM (BY SYSTEM (BY SYSTEM (BY S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 135; DB 1,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                 ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 3010 AA.
                                                                                                                                                                                                                                              CHARGE RELAY S
CHARGE RELAY S
CHARGE RELAY S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             86.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similairy
hes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                               733
1010
1619
1866
2017
3033
3033
3033
1087
1111
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
11169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
1169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=31642;
                                                                                                                                                                                                                                                                 1111
1169
1234
                                                                                                                                                                                                                                347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLG_HCVJT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                  TRANSMEM
                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                  CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                 CHAIN
                                                                                   CHAIN
                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
CARBOHYD
                                               NP_BIND
SITE
                                                              CARBOHYD
                                                                                      CARBOHYD
                                                                                                              CARBOHYD
                                                                                                                      CARBOHYD
                                                                                                                                                                      CARBOHYD
                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                Query Match
              FRANSMEM
                                                                                                                                                                                                                                                       Best Loc
Matches
                                                                                                                                                                                                                                                                                                                      RESULT 12
                                                                                                                                                                                                                                                                                                                               POLG_HCV1
                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                  between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CAPSID PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEIN NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
(RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
                                                                                                                                          SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                 EMBL; D11168; BAA01943.1; --
PIR; A45573; A45573.
PDB; 1AA0; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
MEROPS; S29,001; --
MEROPS; U39.001; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS4a; 1
HCV_NS4b; 1
HCV_NS5a; 1
                                                                                                                                                                                                                                                                            InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                    PF01538; HCV_NS2; 1
PF02907; HCV_NS3; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    729
1006
1615
1862
2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01006; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          192
384
730
1007
1616
                                                                                                                                                                                                                                                                                                                                                                                                           PF01560;
                                                                                                                                                                                                                                                                                                                                                                                                                   PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P55); Nonstructural protein NS5A (P65); NONSTRUCTURAL NS5A (P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
                                                                        SIMILARITY).
                                                                                                                                                                                                                                                             (POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 134; DB 1; Length Ju. Pred. No. 1.2e-09; 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    94A1C77435D642BB CRC64;
                                            CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                  ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               326573 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              85.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    llarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
hes 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus.
NCBI_TaxID=11104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RNA](N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCV1
P26664;
                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                            ACT_SITE
ACT_SITE
ACT_SITE
```

Gaps

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RT A Chen P.J., Lin M.H., Tai K.F., Lin P.C., Lin C.J., Chen D.S.;
RA Chen P.J., Lin M.H., Tai K.F., Lin P.C., Lin C.J., Chen D.S.;
RY The Taiwanese hepatitis C virus genome: sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA.";
R. Virology 188:102-113(1992).
C -1- CHOOTON: THE SMALL PROFIENS NSZB, NSZB, NSZB, AND NSZB ANE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NSJ AND NSS MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 postition, Cys or Thr in Pl and Ser or Ala in Pl'.
C -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNB)(N).
C -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCARSID OF SUBUNIS: PROTEIN E NUCLEOCARSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 42, Last annotation update)
Genome polyprotein [Contains capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NRJ); Protein P1 Nording Protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS58 (P56); Nonstructural protein NS58 (P56); Nonstructural protein NS58 (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (1900ate Taiwan) (HCV).
                                                  (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 131; DB 1; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                      327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                                                GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M84754; -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
  2200454488
64488
62467
62467
78460
78660
78660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCVTW
P29846;
                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                  CARBOHYD
                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       fatches
     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DATE THE PROPERTY OF THE PROPE
                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation use Burpean Bioinformatics Institute. There are no restrictions on its modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELODE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                      SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC. .
N-LINKED (GLCNAC. .
N-LINKED (GLCNAC. .
N-LINKED (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR001319; HCV_NS3.
InterPro; IPR001409; HCV_NS3.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001509; HCV_Cappid; IPP6an; PF01543; HCV_Cappid; IPP6an; PF01543; HCV_Cappid; IPP6an; PF01543; HCV_Cappid; IPP6an; PF01539; HCV_NS1; IPP6an; PF01509; HCV_NS3; IPP6an; PF01001; HCV_NS4b; IPP6an; PF01001; HCV_NS5a; IPP6an; PF01001; HCV_NS1a; IPP6an; PF01001; HCV_N
                                                                                                                                                                                                                                                                                                                                                                              EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PIR; A39166; GNWVC3.
PDB; 1A1V; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
MENOPS; 259. 001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
```

```
MEDINE-98133924; PubMed-9467002;
Chakrabarti L., Bristulf J., Foss G.S., Davies K.E.;
"Expression of the murine homologue of FMR2 in mouse brain and during
                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                  FWR2_HUWAN STANDARD; PRT; 1311 AA.
P51815; 043786; 060215; P78407; 013521; 014323; 09UNA5;
01-007-1996 (Rel. 34, Created)
16-007-2001 (Rel. 40, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Fragila X mental retardation 2 protein (Protein FMR-2) (FMR2P) (0x19 Protein) (Fragile X E mental retardation syndrome protein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gu Y. Shen Y., Gibbs R.A., Nelson D.L.;
"Identification of FMR2, a novel gene associated with the FRAXE CCG
repeat and CpG island.";
Nat. Genet. 13:109-113(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. (ISOFORM 1).
MEDLINE-97446139; PubMed-9299237;
Gecz J., Blelby S., Sutherland G.R., Mulley J.C.;
"Gene structure and subcellular localization of FMR2, a member of new family of putative transcription activators.";
Genomics 44:201-213(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE=Fetal brain;
MEDLINE=96422267; PubMed=8824884;
Chakrabarti L., Knight S.J.L., Flannery A.V., Davies K.E.;
"A candidate gene for mild mental handicap at the FRAXE fragile
                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TISSUE-Fetal brain, and Placenta;
MEDLINE-96241583; PubMed-8673085;
Gecz J., Gedeon A.K., Sutherland G.R., Mulley J.C.;
"Identification of the gene FMR2, associated with FRAXE mental
                                                                 Length 3010;
327047 MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wang L., Thibodeau S.N.;
Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases
                                                           Score 128; DB 1;
Pred. No. 7.6e-09;
1; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 520-1311 FROM N.A. (ISOFORM 3).
Dugan-Rocha S.P., Glbbs R.A.,
Submitted (FEB-1997) to the EMBL/GenBank/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 348-421 FROM N.A. (ISOFORM 1).
                                                                                                                                                                            2 KPOROPKRNTPRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                               TISSUE-Brain;
MEDLINE-96241584; PubMed-8673086;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hum. Mol. Genet. 5:275-282(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hum. Mol. Genet. 7:441-448(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. (ISOFORM 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. (ISOFORM 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. (ISOFORM 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. (ISOFORM 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     retardation.";
Nat. Genet. 13:105-108(1996).
                                                           82.1%;
88.9%;
                                                                                    Local Similarity 88.9
3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE-Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          development.
SEQUENCE
                                                        Query Match
                                                                                                                                                                                                                                                                                                                             RESULT 14
FMR2_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               site."
                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                           DDT TO DE TO
                                                                                                                                                                               ò
                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CELLULAR AMINOPEPTIDASE.

CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOBE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ (POTENTIAL).
PROTENEL PROTEIN NSJ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SMART; SM00487; DEXDC: 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ( POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
CHARGE RELAY SYSTEM (CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                   InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002511; HCV_core.
InterPro; IPR002511; HCV_core.
InterPro; IPR002511; HCV_NS1.
InterPro; IPR001410; HCV_NS2.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR00166; HCV_NS5a.
InterPro; IPR00166; HCV_NS5a.
InterPro; IPR00166; HCV_NS5a.
InterPro; IPR001095; RNA_Pol_DS_PS.
InterPro; IPR0006; HCV_NS2; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS4; I.
Pfam; PF00271; helicass_C; I.
Pfam; PF00271; helicass_C; I.
Pfam; PF00271; helicass_C; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LINKED
LINKED
LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINKED
                                                                     MEROPS; S29.001; --
MEROPS; U39.001; --
InterPro; IPR001410; Di
InterPro; IPR002522; HG
InterPro; IPR002521; HG
         1N64; 25-FEB-03
1N53; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11165
11230
11316
1230
2030
2033
2034
4417
4423
5322
5325
6233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3D-structure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
```

88888888888888888888888

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                             entities requires a license agreement (See http://www.lsb-slb.ch/announce/or send an email to license@lsb-slb.ch),
                                                                                                     IsoId-P51816-3; Sequence-VSP_000211, VSP_000212, VSP_000216;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Missing (in isoform 2, isoform 3 and isoform 4).
/FIId=VSP_000211.
Missing (in isoform 2, isoform 3 and isoform 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     isotorm 4).
/FIId=VSP_000212.
Missing (in isoform 2 and isoform 4).
/FIId=VSP_000213.
                                                                        IsoId=P51816-2; Sequence=VSP_000211, VSP_000212, VSP_000213;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Triplet repeat expansion; Disease mutation; Alternative splicing.
VARSPLIC 57 60 Missing (in isoform ? isoform ?
Event=Alternative splicing; Named isoforms=4;
Comment=Additional isoforms seem to exist;
                                            IsoId-P51816-1; Sequence-Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GO; GO:0007275; P:development; TAS. GO; GO:0009405; P:pathogenesis; TAS. Pfam; PF05110; AF-4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JOINED.
JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AF139977; AAD45878.1; JOINED. AF139978; AAD45878.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOINED,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                         U48436; AAC82513.1; -. L76569; AAA99416.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AF012624; AAB71534.1;
AF012603; AAB71534.1;
AF012604; AAB71534.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB71534.1;
AAB71534.1;
AAB71534.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CAA04822.1
AAB69976.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB71534.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB71534.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB71534.1
AAB71534.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB71534.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB71534.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAD45878.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB71534.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         X95463; CAA64730.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HGNC:3776; FMR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AF012617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AF012610;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AF012615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AF012607
                                                                                        Name=3;
                                                            Name-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30954
                              Name-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Genew;
MIM; 30
                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                 ö
                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                               01-OCT-1996 (Rel. 34, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Growth/differentiation factor 6 precursor (GDF-6) (Cartilage-derived morphogenetic protein 2) (CDMP-2) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                       Bos taurus (Bovine).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GROWTH/DIFFERENTIATION FACTOR 6. BY SIMILARITY.
                                                                                                                         Score 65; DB 1; Length 1311; Pred. No. 0.73;
                                          Missing (in isoform 3).
/FTId-VSP_000216.
D -> A (IN REF. 3).
A -> V (IN REF. 2).
P -> Q (IN REF. 5 AND 6).
MW; 9F151EC06BB636CB CRC64;
                                                                                                                                                9; Indels
N -> K (in isoform 4).
/FTId=vSP_000214.
Missing (in isoform 4).
/FTId=vSP_000215.
                                                                                                                                                                                                                                                   436 AA
                                                                                                                                                7; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PIR: B55422, B55452.
HSSP: P18075; IBMP.
InterPro: IPR0012400; GF_cysknot.
InterPro: IPR00111; TGFb.
InterPro: IPR00111; TGFb.N.
Pfam; PF00049; TGF-beta, 1.
Pfam; PF00049; TGF-beta, 1.
PRNITS; PR00438; GFCyskNoT.
PRODOM; PD000357; TGFb; 1.
PROSTIE; PS00250; TGFB; 1.
GCOWth factor; Cytokine; Glycoprotein.
                                                                                                                                                                                   1 PKPQRQPKRNTPRRPQDVKFPGGGQIV 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                    ₩;
                                                                                                                         41.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; U13661; AAA61416.1; -.
                                                                                                    144739
                                                                                                                                    Local Similarity 40.7 nes 11; Conservative
                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  436
                      1311
                                                                  195
470
548
                                                                                                                                                                                                                                                                                                                                                                            Bovidae; Bovinae; Bos.
NCBI_TaxID-9913;
 466.
                                            971
                                                                                                   1311 AA;
                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
 466
                      467
                                                                                                                                                                                                                                                   GDF6_BOVIN
P55106;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
DISULFID
                                                                              CONFLICT
CONFLICT
SEQUENCE
 VARSPLIC
                                                                                                                           Query Match
                       VARSPLIC
                                              VARSPLIC
                                                                   CONFLICT
                                                                                                                                                                                                                                          GDF6_BOVIN
                                                                                                                                                Matches
                                                                                                                                                                                                                              RESULT 15
                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                            g
```

```
364 433 BY SIMILARITY.
368 435 BY SIMILARITY.
400 400 INTERCHAIN (BY SIMILARITY).
27 N-LIRKED (GLCNAC. ..) (POTENTIAL).
89 89 N-LINKED (GLCNAC. ..) (POTENTIAL).
436 AA; 47873 MW; B0688E12EF8AE91D CRC64;
DISULFID
DISULFID
DISULFID
CARBOHYD
CARBOHYD
SEQUENCE
FT
FT
FT
SO
```

Gaps ò Ouery Match 39.7%; Score 62; DB 1; Length 436; Best Local Similarity 41.7%; Pred. No. 0.56; Matches 10; Conservative 7; Mismatches 7; Indels

ö

q ò

Search completed: August 7, 2003, 11:20:01 Job time : 6.90909 secs

œ

Fri Aug

Perfect score:

Sequence:

protein

š

Run on:

Scoring table:

Searched:

Minimum DB Maximum DB

Database

us-09-491-146a-27.rspt

```
STRAIN-NEL25;

WEDLINE-9420170; PubMed-8151307;

A TOKITA H., Shrestha S.M., Okamoto H., Sakamoto M., Horikita M.,

A Tokita H., Shrestha S.M., Miyakwa Y., Mayumi M.;

A Tokita H., Shrestha S.M., Miyakwa Y., Mayumi M.;

RA Izuka H., Shrestha S.M., Miyakwa Y., Mayumi M.;

RT classification into the third major group.";

RI classification into the third major group.";

RL J. Gen. Virol. 75:931-936(1994)

CC LIPOPROFEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN M AND GLYCOROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF

PROTEIN C AND MRNA (BY SIMILARITY).

DR EMBL; DIGGI4; BAAA04036.1; -.

DR InterPro; IPR002522; HCV_capsid.

DR InterPro; IPR002521; HCV_cope.

DR InterPro; IPR002531; HCV_core.

DR InterPro; IPR002531; HCV_core.

DR InterPro; IPR002531; HCV_core.
                                                                                                                                                                                      hepatitis
hepatitis
                                                                                   hepatitis
hepatitis
                                                                                                              hepatitis
hepatitis
                                                                                                                                                  hepatitis
hepatitis
                    nepatitis
                             nepatitis
                                     hepatitis
                                              hepatitis
                                                       hepatitis
                                                                hepatitis
                                                                          hepatitis
                                                                                                     hepatitis
                                                                                                                                                                    hepatitis
                                                                                                                                                                              hepatitis
                                                                                                                                                                                                          hepatitis
                                                                                                                                                                                                                  hepatitis
                                                                                                                                                                                                                           hepatitis
                                                                                                                                                                                                                                    hepatitis
                                                                                                                                                                                                                                               hepatitis
                                                                                                                                                                                                                                                       nepatitis
                                                                                                                                                                                                                                                                hepatitis
                                                                                                                              hepatitis
                                                                                                                                          hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                           Q8g3j6
Q68883
Q68885
Q81486
Q81542
Q81542
Q81542
Q81542
Q855
                                                                                                                      0893×x0
0893×x0
0893×x0
06892
068796
068796
068796
068798
068798
068798
068798
068798
068798
068798
                                                                                                                                                                                                                                                                                                                                      081548 PRELIMINARY; PRT; 415 AA.
081548;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OT-2002 (TrEMBLrel. 22, Last annotation update)
Core, env, and part of E2/NS1 (Genome polyprotein) (Fragment).
Hepatitis C virus.
                   08qp62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Shrestha S.M., Tsuda F., Okamoto H., Tokita H., Horikita M., Tanaka T., Miyakawa Y., Mayumi M.; Hopatitis B virus subtypes and hepatitis C virus genotypes patients with chronic liver disease in Nepal."; Hepatology 19:805-809(1994).
                                                                                                                                                                                                                                                                                            ALIGNMENTS
                                                              08JP06
081542
08JWK6
                                                                                                           070635
08JWM0
08JWK7
08JWK1
                                                                                                                                                         Q68796
Q81266
Q68966
P88803
                                                                                                                                                                                              Q9DIT6
Q68798
Q68305
Q81807
Q68951
                                            Q68885
Q81486
                                                                                         Q8JWK0
Q8V7J4
                                                                                                                                                  068126
                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-94186155; PubMed-8138250;
                                                                                                                                                                                                                  43
109
782
802
3033
88
                                                                                                                                                                            640
3010
3011
3022
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus.
NCBI_TaxID=11103;
                  STRAIN-NE125
RESULT 1
081548
Q81548 hepatitis c
Q8bowl hepatitis c
Q81230 hepatitis c
Q81231 hepatitis c
Q81231 hepatitis c
Q81231 hepatitis c
Q81234 hepatitis c
Q81236 hepatitis c
Q68861 hepatitis c
Q68861 hepatitis c
Q68865 hepatitis c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       hepatitis
hepatitis
                                                            August 7, 2003, 11:05:41; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q8v7t0
Q8v7s4
                                                                                                                                                                                     830525
                 Compugen Ltd.
        5.1.6
                                                                                                                                                                                   tal number of hits satisfying chosen parameters:
                                                                                                   US-09-491-146A-27
156
1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                   830525 segs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUMMARIES
       GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     081548
08BCW3
081230
081235
081231
081231
081232
081234
081236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q68863
Q68865
Q8V7S2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q8V7T0
Q8V7S4
                                                                                                                                                                                                                                                                                                                                                                                                       sp_unclassified:*
                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.
                                                                                                                                                                                                                                                                                                           sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                      sp_virus:*
sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                               sp_organelle:*
sp_phage:*
                                                                                                                                                                                                        seq length: 0
seq length: 2000000000
                                                                                                                                                                                                                                                                                 sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                   sp_plant:*
sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  %
Query
Match Length DB
                                                                                                                                                                                                                                                                        SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                   sp_fung1:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11009
11009
11009
11009
11009
11000
11001
11001
                                                                                                                                                                                                                                                                                                                                          sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100.0
92.3
91.7
91.7
91.7
91.7
91.7
91.7
91.7
```

Score

Result Š 

```
109 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE 109 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=11103;
                                                                                                                                                                                                                      109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-Ban2-1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-Ban2-3;
                                                                                                                                                                                      Polyprotein.
NON_TER 1
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             081235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        081233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 235 281 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SPERMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                            ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                            Gaps
Pfam; PF01542; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01560; HCV_env; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein; Polyprotein; Transmembrane.
NON_TER 415 415 415 SEQUENCE 415 AA; 44880 MM. OOANN_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ohnor T., Mizokami M., Chstopher T., Ohba K., Suzuki K., Roger W.;
"The nucleotide sequence of the core region of HCV in Pakistan and
Bangladesh: The geographic characterization of HCV in South Asia.";
Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
-i- SUBUNIT: THE 91410 OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Theamboonlers A., Bedl K., Chinchal T., Sriponthong M., Chantarasamee P., Poovorawan Y.; Molecular characterization of HCV core region of HCV infection in Thai blood donor."; Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                           100.0%; Score 156; DB 12; Length 415; 100.0%; Pred. No. 1.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 144; DB 12; Length 119;
Pred. No. 1.9e-12;
1; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              119 119 AA; 13427 MW; 0EBB321B6CB26AOA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     081230 PRELIMINARY; PRT; 109 AA.
081230; CTEMBLrel. 01, Created)
01-NOV-1996 (TEMBLrel. 01, Last sequence update)
01-NOV-1996 (TEMBLrel. 22, Last annotation update)
01-OCT-2002 (TEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAR-2003 (TrEMBLrel. 23, 01-MAR-2003 (TrEMBLrel. 23, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.3%;
ilarity 92.9%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-061653a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-Ban1-1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              081230
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
20120
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OBBCW3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                             SO TE SO ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Othor T., Mizokami M., Chstopher T., Ohba K., Suzuki K., Roger W.;
"The nucleotide sequence of the core region of HCV in Pakistan and Bangladesh: The geographic characterization of HCV in South Asia.";
Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MANA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; D29647; BA31022.1; -.
InterPro; IPR002522; HCV_capsid.

Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ohno T., Mizokami M., Chstopher T., Ohba K., Suzuki K., Roger W.;
                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                          91.7%; Score 143; DB 12; Length 109; 92.9%; Pred. No. 2.4e-12; 1ve 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91.7%; Score 143; DB 12; Length 109; 92.9%; Pred. No. 2.4e-12; Live 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    109
12235 MW; D5D1BBB10A7E3CAB CRC64;
                                                                                                                                                               109
12293 MW; 524E8D425A7E3CB2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1996 (TTEMBLrel. 01, Created)
01-NOV-1996 (TTEMBLrel. 01, Last sequence update)
01-OCT-2002 (TTEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-0CT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               109 AA.
                                                                                                                                                                                                                                                                                                                                                 1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity 92.9%
                                                                                                                                                                                                                         Query Match
Best Local Similarity 92.99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
```

```
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
nes 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
          NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-Ban2-2;
                                                                                                                                                                                                                                                                                                            Polyprotein.
NON_TER 1
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
NON_TER 1
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q81234
Q81234;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q81236
Q81236;
                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 9
Q81236
ID Q812
AC Q812
DT 01-N
DT 01-N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               081234
          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
Othor T., Mizokami M., Chstopher T., Ohba K., Suzuki K., Roger W.;
"The nuclectide sequence of the core region of HCV in Pakistan and
Bangladesh: The geographic characterization of HCV in South Asia.";
Submitted (MAR.1994) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPMOTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIXCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            081232 PRELIMINARY; PRT; 109 AA.
081232;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 22, Last annotation update)
COCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                   Score 143; DB 12; Length 109;
Pred. No. 2.4e-12;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 143; DB 12; Length 109;
Pred. No. 2.4e-12;
1; Mismatches 1; Indels (
                                                                                                                                                                                                                            109 109
109 AA; 12251 MW; D5D1A6060A7E3CAB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    109 AA; 12261 MW; 461E99060A7E3CA9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                109 AA
                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                              InterPro; IPR002522; HCV_capsid. Pfam: PF01543; HCV_capsid; 1. P01yprotein. 109 109 SEQUENCE 109 AA; 12251 MW; D'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                     91.7%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D29648; BAA21023.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (TrEMBLrel. 01, (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                     Local Similarity
nes 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-Ban1-2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-OCT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                     Query Match
Best Local 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           081231
                                                                                                                                                                                                                                                                                                                          4atches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 7
08123
ID 08123
AC 0812
DT 01-N
DT 01-N
DT CORE
OS HEPA
    RATE CCC CCC FT TO DAY TO DAY THE DAY 
                                                                                                                                                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Soria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91.7%; Score 143; DB 12; Length 109; llarity 92.9%; Pred. No. 2.4e-12; Conservative 1; Mismatches 1; Indels (
                                                                                                                                                                                      Length 109;
                                                                                                                                                                                                              Indels
                                                                                                                                                            461E99060A7E3CA9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DSD1A6060A7E3CAB CRC64;
                                                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                   Score 143; DB 12;
Pred. No. 2.4e-12;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Created)
Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                   Core protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                          109 AA
                                                                                                                                                                                                                                      1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                 1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRQTKRNTPRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                  109 109
109 AA; 12261 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             109 AA; 12251 MW;
                                                                                                                                                                                   91.78;
92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01,
                                                                                                                                                                                                           26; Conservative
                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
```

```
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=42791;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCB1_TaxID=42792;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-IND 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
NON_TER
NON_TER
SEQUENCE 1:
                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus
                                    7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                       Q68861
Q68861;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          068863
                                                                                                         RESULT 11
Q68861
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              068863
                                    a
                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Panigrahi A.K., Roca J., Acharya S.K., Jameel S., Panda S.K.;

"Genctype determination of hepatitis C virus from Northern India:
"Genctype determination of hepatitis C virus from Northern India:
"I dentification of a new subtype.";

"I med. Virol. 48:191-198(1996).

"I Subunit: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPPOPOTEIN B. AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MANA (BY SIMILARITY).

"PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MANA (BY SIMILARITY).

"REMBL: X91300: CAA62674.1; -...

"REMBR: X91300: CAA62674.1; -...

"REMBL: X91300: CAA6267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatliis C virus.
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepaclvirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                   Hepatitis C virus type 3.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91.7%; Score 143; DB 12; Length 109; 92.9%; Pred. No. 2.4e-12; Live 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91.7%; Score 143; DB 12; Length 150; 92.9%; Pred. No. 3.4e-12; ative 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           109 109 MW; AGEDG114583E2EEC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      150 150 150 150 MW; DE75E728FF39B2B3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) Capsid protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            150 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-IND 674;
MEDLINE-96432288; PubMed-8835354;
                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 92.99
Matches 26, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity .92.98
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-40363;
                                                                                                                                                                                                   STRAIN-Ban3-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV CORE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 10
Q68867
```

g

ò

RAN OCCOOR DE PARTICION DE PART

```
A MEDLINE-9642288; PubMed-8835354;
A MEDLINE-9642288; PubMed-8835354;
A Panigrahi A.K., Roca J., Acharya S.K., Jameel S., Panda S.K.;
T "Genotype determination of hepatitis C virus from Northern India:
T "Identification of a new subtype.";
T "Identification of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              K
                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus type 3g.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Panigrah A.K., Roca J., Acharya S.K., Jameel S., Panda S.K.;
"Genotype determination of hepatitis C virus from Northern India:
Identification of a new subtype.";
J. Med. Virol. 48:191-188(1996).
-1- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 150 AA; 16503 MW; 2C38B48C6BB9C16A CRC64;
                                                                                                                (TrEMBLrel. 01, Last sequence update) (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 143; DB 12;
Pred. No. 3.4e-12;
                                                                                                                                                                                        Capsid protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Capsid protein (Genome polyprotein) (Fragment)
    150 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           150 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 PKPQRQTKRNIPRRPQNVKFPGGGQIVG 29
                                                                                Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
STRAIN-IND 1452;
MEDLINE-96432288; PubMed-8835354;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91.7%;
92.9%;
                                                                    01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                 Hepatitis C virus type 3b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26; Conservative
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                150
```

ö

```
Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                          Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
-!- SUBMITT: THE VIRION OF THIS VIRUS IS A NUCLECCARSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GIYCORROTEIN E. THE NUCLECCARSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; AB062202; BAB03396.1; -.
InterPro; IPR002521; HCV_copeid
  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                NON_TER 191 191
SEQUENCE 191 AA; 20740 MW; 64151CA83695EF34 CRC64;
                                                                                                                                                                                                                                                                                                                                      1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                            Pfam; PF01543; HCV_caps1d; 1. Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                     91.78;
92.98;
                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 92.99
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
26; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     rissue-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                TISSUE-Liver;
                                                                                                                                                                                                                                       Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          lesions.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kato N.;
                                                                                                    lesions.
                                                                           Kato N.
                                                                                                                                                                                                                                                                                                                                                                                                                         Q8V7T0
                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                            Panigrahi A.K., Roca J., Acharya S.K., Jameel S., Panda S.K.;
"Genotype determination of hepatitis C virus from Northern India :
Identification of a new subtype.";
J. Med. Virol. 48:191-198(1996).
-: SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY IIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 150;
                                                                                                                          Length 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                 150 AA; 16799 MW; BC2932134026E5FD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16728 MW; BC2D761707FCE1B9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-WAR-2002 (TrEMBLrel. 20, Created)
01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
01-CCT-2002 (TrEMBLrel. 22, Last annotation update)
CORE protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Capsid protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 143; DB 12;
Pred. No. 3.4e-12;
1; Mismatches 1;
                                                                                                                          DB 12;
                                                                                                                        91.7%; Score 143; DB 12; 92.9%; Pred. No. 3.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              191 AA.
                                                                                                                                                                                                                                                              150 AA.
                                                                                                                                                 1; Mismatches
                                                                                                                                                                                       2 PKPQRQTKRNTPRRPQNVKFPGGGQIVG 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPOROPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2 PKPQRQTKRNTPRRPQNVKFPGGGQIVG 29
                                                                                                                                                                        1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-96432288; PubMed-8835354;
EMBL; X91306; CAA62680.1; -.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       91.78;
92.98;
                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus type 3g.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 92.9°
Matches 26; Conservative
                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    150
                                                                                       150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1
150 1
150 AA;
                                                                                                                                     Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=42792;
                                                                          150
                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-IND 1751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein.
NON_TER
NON_TER
SEQUENCE 15
                                                              Polyprotein.
NON_TER
NON_TER
SEQUENCE 15
                                                                                                                          Query Match
                                                                                                                                                                                                                                                                        068865;
                                                                                                                                                                                                                                                            068865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q8V7S2
                                                                                                                                                                                                                                     TRULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
                                                                                                                                                                                                                                                3865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DT DT DT OS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FTS
  ç
                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                        ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

Gaps

ö

Length 191; 1; Indels

Score 143; DB 12; Pred, No. 4.4e-12;

1; Mismatches

BY A

```
Hepatitis C virus guasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIFOPPROTRIN BYNELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX.OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX.OF EMBL. AB063194; BAB83388.1; ...
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_corps.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_corps.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191 AA; 20740 MW; 64151CA83695EF34 CRC64;
                                                                                                                                                                                                                      U1-WAR-2002 (TrEMBLrel. 20, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 143; DB 12;
Pred. No. 4.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Mismatches
                                                                                                                                                        191
1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAR-2002 (TrEMBLrel. 20, Created)
01-WAR-2002 (TrEMBLrel. 20, Last seq
01-0CT-2002 (TrEMBLrel. 22, Last ann
                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91.78;
92.98;
```

Search completed: August 7, 2003, 11:19:00 Job time: 26.6364 secs

1, Appli 2, Appli 2, Appli 16, Appl 1, Appli

Sequence Sequence

App App App App App App App

Sequence Sequence Sequence Sequence

US-08-836-075A-46 US-08-810-195-1 US-08-601-195-1 US-08-61-1352-2 US-07-681-701-16 US-08-290-65A-155 US-08-290-65A-155 US-08-290-65A-157 US-08-290-65A-157 US-08-290-65A-167 US-08-290-65A-161 US-08-290-65A-161 US-08-290-65A-161 US-08-290-65A-161 US-08-290-65A-161 US-08-290-65A-163 US-08-290-65A-163 US-08-290-65A-164 US-08-290-65A-164 US-08-290-65A-164

159, 160, 161, 163, 164,

Sequence Sequence Sequence Sequence Sequence 3 Sequence

ALIGNMENTS

```
GENERAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAI
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATEMIN Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match 100.0%; Score 156; DB 3; Best Local Similarity 100.0%; Pred. No. 4.6e-14; Matches 28; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATORNEY/AGENT INFORMATION:
NAME: WARREN, WILLTAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
TELECPHONE: 404-818-3709
TELEPHONE: 404-818-3799
                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 27, Application US/08921887 Patent No. 6030771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEFAX: 404-B18-5/52
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
TENGTH: 28 aming acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY: not reach
MOLECULE TYPE: protein
COMPUTER READABLE FORM: MEDIUM TYPE: Floppy of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CIT:. GASTATE: GACOUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FRAGMENT TYPE:
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                       RESULT 1
US-08-921-887-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-921-887-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      App11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 27, Appl
                                                                                                                            (without alignments)
112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 175,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 3,
Sequence 10,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 26,
                                                                                                          August 7, 2003, 11:07:41; Search time 10.5455 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sednence
Sednence
Sednence
Sednence
Sednence
Sednence
Sednence
Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                       328717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
/cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
/cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
/cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
/cgn2_6/ptodata/1/1aa/PcTUS_COMB.pep:*
/cgn2_6/ptodata/1/1aa/PcTUS_COMB.pep:*
             5.1.6
Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-836-075A-104
US-08-921-887-26
US-08-290-665A-175
PCT-US95-10398-175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-836-075A-10
US-08-635-886C-198
US-08-635-886C-232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-389-756-1
US-08-380-160-1
US-08-380-160-1
US-08-083-947-23
US-08-262-033-08-262-03
US-08-262-037-26
                                                                                                                                                                                                                                                                                                                  al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PCT-US94-07088-23
                                                                                                                                                                        US-09-491-146A-27
156
1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-380-160-6
US-09-020-846-36
US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-921-887-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PCT-US95-13660-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -08-904-686A-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              us-08-501-195-2
                                                                                                                                                                                                                                                                                    328717 seqs, 42310858 residues
              GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                       BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                   Minimum DB seq length: 0 Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    249
119
119
144
144
144
144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.℃ 0
.. ..
                                                                                                                                                                        Title:
Perfect score:
                                                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score
                                                                            OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Database :
                                                                                                                                                                                                           Sequence:
                                                                                                                                                                                                                                                                                       Searched:
                                                                                                            Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Result
No.
```

Gaps

ö

Length 28; Indels

```
Sequence 26, Application US/08921887

Sequence 26, NOWARD A.

TILE OF INVENTION: MOSALC PROTEIN AND RESTRICTION

TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME

NUMBER OF SEQUENCES: 55

CORRESPONDENCE ADDRESS:

ADDRESSEE: JONES & ASKEW, LLP

STREET: 191 Peachtree Street, N.W., 37th Floor

CITY: Atlanta

STATE: GA

COUNTRY: USA

STATE: GA

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: TEM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARR: PAPLICATION DATA: GO, ADDRESSEE

ADDRESSEE: Patentin Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 89.1%; Score 139; DB 3; Length 74; Best Local Similarity 92.9%; Pred. No. 2.1e-11; Matches 26; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                   STATE: IGARD
COUNTRY: USA
ZIP: 77210-4433
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBN PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft Word 6.0 / ASCII tex
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,075A
FILING DATE: 21 Apr 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/04155
FILING DATE: 23 Oct 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORREY/ABENT INFORMATION:
NAME: RAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 104:
FENGTHY: 74 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: US/08/921,887
                 ADDRESSEE: ARNOLD, WHITE STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION: NAME: WARREN, WILLIAM L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear
MOLECULE TYPE: peptide
CORRESPONDENCE ADDRESS: ADDRESSE: ARNOLD, W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid
                                                                                                 TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-836-075A-104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-921-887-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                       APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 104, Application US/08836075A
Patent No. 6180768
GENERAL INFORMATION:
APPLICANT: MARRIENS, GERFT
APPLICANT: STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: ADD THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC;
TITLE OF INVENTION: AGENTS
NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 89.1%; Score 139; DB 3; Length 28; Best Local Similarity 92.9%; Pred. No. 7.6e-12; Matches 26; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                 1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              03063-0380
                                                                                                                                                                                Sequence 24, Application US/08921887
Patent No. 6030771
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE/DOCKET NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 0306:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3799
INFORMATION FOR SEC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY: not relevant MOLECULE TYPE: protein HYPOTHETICAL: NO ANTI-SPACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HATI-SENSE: NO FRAGMENT TYPE: internal ORIGINAL SOURCE: ORGANISM: Hepatitis virus US-08-921-887-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: amino acid
STRANDEDNESS: not
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 3
US-08-836-075A-104
                                                                                                                                                                 US-08-921-887-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STATE:
                    ô
                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 175, Application PC/TUS9510398
GENERAL INFORMATION:
APPLICANT: BURKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: CORE GENES OF SIGOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AMIN ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 191;
                                                                                       Score 136; DB 2; Length 191;
Pred. No. 1.4e-10;
                                                                                                                                                         2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 136; DB 5; Length 19
Pred. No. 1.4e-10;
1; Mismatches 2; Indels
                                                                 87.2%; Scc. 89.3%; Pred. No. 1. Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPORKTKRNTSRRPODVKFPGGGQIVG 32
                                                                                                                                                                                                              1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                   5 PKPQRKTKRNTSRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 205
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              175:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          87.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: WORDPERFECT 5.1 CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US
FILING DATE: 15-AUG-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (212) 758-4800
(212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 191 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STREET: 345 PARK AVENUE
CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INDIVIDUAL ISOLATE: P8 PCT-US95-10398-175
; INDIVIDUAL ISOLATE: P8
US-08-290-665A-175
                                                                                       Query Match
Best Local Similarity 89.33
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEX: 421792
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGINAL SOURCE
                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
PCT-US95-10398-175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUCLEOTIDE AND DEDUCED
ANINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 3; Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match 88.5%; Score 138; DB 3; Best Local Similarity 89.3%; Pred. No. 1e-11; Matches 25; Conservative 1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPORKPNRNTNRRPODVKFPGGGOIVG 28
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3799
TELEFAX: 404-818-3799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US/08/290,665A FILING DATE: 15-AUG-1994 CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 175, Application US/08290665A Patent No. 5882852 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
                                                                                                                INFORMATION FOR SEQ ID NO: 26: SEQUENCE CHARACTERISTICS: LENGTH: 28 amino acids TYPE: amino acid STRANBENESS: not relevant control of the con
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATTORNEY/ACENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 20,7
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 175:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                      STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                            FRAGMENT TYPE: internal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE OF INVENTION: NUCL
TITLE OF INVENTION: AMI
TITLE OF INVENTION: AMI
TITLE OF INVENTION: SEQ
NUMBER OF SEQUENCES: 26.
CORRESPONDENCE SEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SS: unknown unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  homosapiens
                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: amino acid
STRANDEDNESS: unh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                   <u>2</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-290-665A-175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; ORGANISM:
US-08-921-887-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                      ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

us-09-491-146a-27.rai

```
APPLICANT: DALBON, PASCAL
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY
TITLE OF INVENTION: FOR DETECTING THE LATTER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: BM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,846
FILNG DATE: 09-FEB-1998
CLASSIFICATION WIMBER: US 9-027015
FILNG DATE: 10-FEB-1997
PRIOR APPLICATION NUMBER: JP 9-027015
FILING DATE: 10-FEB-1997
PRIOR APPLICATION NUMBER: JP 8-024045
FILING DATE: 09-FEB-1996
ATPORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SYSTEM: PC-DOS/MS-DOS
Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 135; DB 4;
Pred. No. 3.9e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPORQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPORKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 2, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: Wegner, Harold C. REGISTRATION NUMBER: 25,258 REFERENCE/DOCKET NUMBER: 05. TELECOMMUNICATION: TELEPHONE: (202) 672-5300
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: OLIFF & BERRIDGE STREET: P.O. BOX 19928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CURRENT APPLICATION DATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; MOLECULE TYPE: peptide US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                               STREET: 3000 K Str
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 9
US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                            AND WHICH CAN BE USED ESPECIALLY
                                                                                                                                                               SYNTHETIC POLYPEPTIDES BELONGING TO THE HEPATITIS C VIRUS (HCV) AND WHICH CAN BE FOR DETECTING THE LATTER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 135; DB 3; Length 34;
Pred. No. 3.1e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READBALE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHIMERA ANTIGEN PEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/POCKET WMBER: WPB 28682
TELECOMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human Hepatitis C Virus
                                                    Sequence 6, Application US/08380160
Fatent No. 6235284
GENERAL INFORMATION:
APPLICANT: DALBON, Pascal
APPLICANT: JOLLVET, MICHEL
TITLE OF INVENTION: HEPATITIS C VI
TITLE OF INVENTION: HEPATITIS C VI
TITLE OF INVENTION: FOR DETECTING
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: YANAGUCHI, Kenjiro
APPLICANT: KASHIWAKUMA, TOMIKO
APPLICANT: CHIBA, YUKIE
APPLICANT: YAGI, Shintaro
APPLICANT: HASEGAWA, Akira
TITLE OF INVENTION: CHIMERA ANTIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-020-846-36
; Sequence 36, Application US/09020846
; Patent No. 6322965
; GENERAL INFORMATION:
                                                                                                                                                                                                                                              STATE: VA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FRAGMENT TYPE: N-terminal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOLECULE TYPE: peptide HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM:
US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                              COUNTRY:
                                  US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
Michel
SYMMETIC POLYPEPTIDES BELONGING TO THE
HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALL)
FOR DETECTING THE LATTER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                               1; Mismatches
                                                                                                                                    Score 135;
Pred. No. 46
                                                                                                                                                                                                                           1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                  4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-WAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Human Hepatitis C Virus STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                           Sequence 1, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: Berridge, William P. REGISTATION NUMBER: 30,024 REFERENCE/DOCKET NUMBER: WPB TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BERRIDGE
                                                                                                                                    86.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: DALBON, PASCAL
APPLICANT: JOLIVET, MICHEL
TITLE OF INVENTION: SYNTHE
TITLE OF INVENTION: FOR DE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRI
STREET: P.O. BOX 19928
CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
  ; VOLUME: 88
; PAGES: 3392-3396
; DATE: 1991
; RELEVANT RESIDUES: 2 TO 45
US-09-389-756-1
                                                                                                                                    Query Match
Best Local Similarity 89.33
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COTHER INFORMATION: /
COTHER INFORMATION: p
COTHER INFORMATION: t
US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MOLECULE TYPE: peptide HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY: Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HYPOTHETICAL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
                                                                                                                                                                                                                                                                                                                                    RESULT 11
US-08-380-160-1
                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION:
APPLICANT: DOLIVET, MICHEL
APPLICANT: DOLIVET, MICHEL
APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: LACONX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: TREATING AN HCV INFECTION
TITLE OF INVENTION: TREATING AN HCV INFECTION
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILE REFERENCE: 103959
CURRENT APPLICATION NUMBER: DS/09/389,756
CURRENT FILING DATE: 1999-09-07
EARLIER FILING DATE: 1999-03-05
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PATCHILIN VOY: 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
tent No. 6576240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 86.5%; Score 135; DB 3; Length 44; Best Local Similarity 89.3%; Pred. No. 4e-11; Matches 25; Conservative 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPOROPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                             NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ANTI-SENSE: NO FRAGMENT TYPE: N-terminal ORIGINAL SOURCE: ORGANISH: Human Hepatitis C Virus STRAIN: H77
FILING DATE: CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-YAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 1, Application US/09389756 Patent No. 6576240
                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY: Ilnear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME/KET: Peptide
COCATION: 1..44
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 10
US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQ ID NO 1
LENGTH: 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               atent No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Hosein, Barbara
APPLICANT: Wang, Chang Yi
TILLE OF INVENTION: Peptides Effective for Diagnosis and
TITLE OF INVENTION: Detection of Hepatitis c Infection
                                                                                                                                             COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPOROPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 PKPQKKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                         CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,947
FILING DATE: 19930628
                                                                                                                                                                                                                                               FILING DATE: 19930628
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 946,054
FILING DATE: 15-SEP-1993
ATTORNEY/AGGNI INCRMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US/08/530,550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 3, Application US/08530550 Patent No. 5736321 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE/DOCKET NUMBER: 20
TELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    86.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                           (516)273-2828
(516)273-1717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,
REFERENCE/DOCKET NUMBER:
          25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE: (516)273-2828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO:
SEQUENCE CRARACTERISTICS:
LENGTH: 61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 89,39
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; MOLECULE TYPE: peptide US-08-083-947-23
                                                                                                             COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
STREET: 25 bu.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hauppauge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       USA
                                                                                     11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEPHONE:
                                                                      COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 23, Application US/08083947
Patent No. 5639594
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
APPLICANT: Hosein, Barbara
TITLE OF INVENTION: No. 5639594el Linear and Branched Peptides Effective
TITLE OF INVENTION: in Diagnosing and Detecting No. 5639594-B Hepat
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: M. Lisa Wilson
                                                                                                                                                                                                                                                                             GENERAL INFORMATION:

APPLICANT: Wang, Chang Yi

TILLE OF INVENTION: No. 5582968-1 Branched Hybrid and Cluster

TITLE OF INVENTION: No. 5582968-B Hepatitis

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:
ADDRESSEE: United Biomedical Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                  ;
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .;
0
                   86.5%; Score 135; DB 3; Length 45; 89.3%; Pred. No. 4.1e-11; 1ve 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 135; DB 1; Length 61;
Pred. No. 5.7e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/946,054
FILING DATE: 1.5.SEP-1992
                                                                                                  1 PKPOROPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                            5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 E: United Biomedical Inc.
25 Davids Dr.
                                                                                                                                                                                                                      US-07-946-054-9; Sequence 9, Application US/07946054; Patent No. 5582968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATTORNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/CDCKET NUMBER: 2000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-273-2828
TELEFAX: 516-273-1717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX: 516-273-1717 INFORMATION FOR SEQ ID NO: 9:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS: LENGTH: 61 amino acids
                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: USA
ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: peptide 07-946-054-9
                   Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hauppauge
New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET:
                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

ô

Gaps

ö

```
ò
                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
APPLICART: Chang YI Wang and Barbara Hosein
TITLE OF INVENTION: SYNTHETIC PEPTIDES SPECIFIC FOR
TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
TITLE OF INVENTION: INFECTION AND PREVENTION THEREOF AS VACCINES
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                 Score 135; DB 1; Length 61;
Pred. No. 5.7e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 135; DB 1; Length 61; Pred. No. 5.7e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1151-4043 US3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/262,037
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/719,819
FILING DATE: 24-June-1991
APPLICATION NUMBER: 07/651,735
FILING DATE: 11-Mar-1991
APPLICATION NUMBER: 07/651,735
FILING DATE: 07-Feb-1991
APPLICATION NUMBER: 07/558,799
FILING DATE: 16-April-1990
APPLICATION NUMBER: 07/510,153
FILING DATE: 16-April-1990
ATORNEY/AGENT INFORMATION:
REGISTRATION NUMBER: 29,323
REGISTRATION NUMBER: 29,323
REGISTRATION NUMBER: 29,323
RELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVE. CITY: NEW YORK STATE: NEW YORK COUNTRY: USA 21P: 10154 COMPUTER READABLE FORM: MEDIUM TYPE: FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 26, Application US/08262037 Patent No. 5747239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                     86.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    86.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 61 amino acids TYPE: Amino acid
TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                 Query Match 86.5'
Best Local Similarity 89.3'
Matches 25; Conservative
                                                                                                                                                                                          TOPOLOGY: Ilnear
; MOLECULE TYPE: peptide
US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 15
.-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

Search completed: August 7, 2003, 11:23:51 Job time: 11.6364 secs

APPI APPII APPII APPII APPI

Sequence Seq

```
COUNTRY: 1000

ZIP: 77210-4433

COMPUTER READABLE: FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC. DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
APPLICATION DATA:
APPLICATION DATA:
FILING DATA:
PRIOR APPLICATION DATA:

FILING DATA:

FILING DATA:

CONNERS APPLICATION DATA:

FILING DATA:

FIL
US-09-878-281-44
US-109-929-955-2
US-10104-966-2
US-09-194-949-3
US-09-851-138-76
US-09-851-138-18
US-09-851-138-18
US-09-899-046-52
US-09-899-046-54
US-09-878-281-52
US-09-878-281-52
US-09-973-025-50
US-09-973-025-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
                                                                                                                                                                                                                                                                                                                                                                                    US-09-742-659-4

US-09-952-572-9

US-09-929-955-1

US-09-747-419-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 104, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
STUTVER, LIEVER
TITLE OF INVENTION: NEW SEQUENCES OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                        10
     INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                     2894
2985
3011
3011
3011
3011
3011
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STATE: TEXAS
 US-09-851-138-104
   Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 7, Appli
Sequence 71, Appl
Sequence 14, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 152, Appl
Sequence 152, Appl
Sequence 44, Appl
Sequence 44, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 104, App
                                                                                                                    August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         rulismed. Applications_Ash.

(gn2_6/ptodata1/pubpaa/USO7_PUBCOMB.pep:*
(gn2_6/ptodata1/pubpaa/USO7_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO6_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO6_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO6_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_PUBCOMB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_PUBCOMB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_PUBCOMB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_PUBCOMB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_PUBCOMB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_PUBCOMB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_PUBCOMB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_NEW_PUB.pep:*
(gn2_6/ptodata1/pubpaa/USO8_NEW_PUB.pep:*
                                                                                                                                                                                                                                                                                                                                                       451899
                   GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-851-138-104
US-10-367-677-1
US-09-851-138-10
US-09-756-875-8
US-09-756-875-8
US-09-921-397-77
US-09-921-397-78
US-09-921-397-78
US-09-921-397-78
US-09-921-397-78
US-09-921-397-78
US-09-921-397-78
US-09-931-138-46
US-09-931-138-46
US-09-931-138-46
US-09-931-138-60
US-09-931-138-60
US-09-978-281-152
US-09-978-281-152
US-09-978-281-152
                                                                                                                                                                                                                                                                                                                                                       al number of hits satisfying chosen parameters:
                                                                                                                                                                                                               156
1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                    451899 seqs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                       protein search, using sw model
                                                                                                                                                                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                         seq length: 0
seq length: 200000000
                                                                                                                                                                                              US-09-491-146A-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            44
74
91
97
103
1133
1137
1166
1169
1169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          135
135
135
135
135
135
135
135
135
135
                                                                                          protein
                                                                                                                                                                                                                                                                                                                                                                                        Minimum DB
Maximum DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Database :
                                                                                                                                                                                                                               Sequence:
                                                                                                                                                                                                                                                                                                                      Searched:
                                                                                                                         Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
```

HEPATITIS C VIRUS GENOTYPES PROPHYLACTIC, THERAPEUTIC AND DIAGN

DURKEE

Sequence 23, Appl Sequence 40, Appl Sequence 4, Appli Sequence 9, Appli Sequence 1, Appli Sequence 20, Appli

3, Appli 1, Appli 20, Appl 2, Appli

Sequence 3 Sequence 1 Sequence 2

Sequence

us-09-491-146a-27.rapb

ò

```
Sequence 1, Application US/09758308

Patent No. USZ0020090607A1

GENERAL INFORMATION.

GENERAL INFORMATION.

TITLE OF INVENTION: PROTEINS

TITLE OF INVENTION: PROTEINS

FILE SEFERBNCE: 14114.034942

CURRENT REFERENCE: 14114.034942

CURRENT APPLICATION NUMBER: US/09/758,308

CURRENT FILING DATE: 2001-01-10

PORTOR APPLICATION NUMBER: 05.001-01-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM FC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICTOSOft Word 6.0 / ASCII text output
CURRENT APPLICATION NDATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: B 94870166.9
FILING DATE: CUNKNOWN>
APPLICATION NUMBER: B 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 22 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match

86.5%; Score 135; DB 10;
Best Local Similarity 89.3%; Pred, No. 5.7e-10;
Matches 25; Conservative 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 135; DB 9;
Pred. No. 7.1e-10;
                                                    ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MBER: 60/092,339
1999-07-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 74 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION
                                                                             STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PatentIn version 3.0
         NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86.5%;
ilarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: Hepatitis C Virus US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEC ID NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRIOR APPLICATION NUMBER:
PRIOR FILING DATE: 1999-0
NUMBER OF SEQ ID NOS: 5
                                                                                             CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOFTWARE:
SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 10, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAETENS, GEERT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: JOLIVE, MICHEL
APPLICANT: JOLIVE, MICHEL
APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILE REFERENCE: 103999
CURRENT APPLICATION NUMBER: US/10/367,677
CURRENT FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: US/09/389,756
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/FR98/00442
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PRIOR PLING DATE: 11999-09-07.
PRIOR PILING DATE: EARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SECTION OF THE STANDARD NOS: 11
                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
JULE: of Hepatilis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
                                                                                                                                                                                Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                         89.1%; Score 139; DB 10;
llarity 92.9%; Pred. No. 1.8e-10;
Conservative 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 135; DB 15;
Pred. No. 3.4e-10;
LENGTH: 74 amino acids
TYPE: amino acid
TYPE: amino acid
TOPOLOGY: 11near
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 104:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Mismatches
                                                                                                                                                                                                                                                                1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                       4 PKPQRKTKRNINRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 1, Application US/10367677 ublication No. US20030118604A1 ENEAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               86.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Hepatitis C virus PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RELEVANT RESIDUES: 2 TO 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 89.39
Matches 25; Conservative
                                                                                                                                                                         Ouery Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PAGES: 3392-3396
DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 3
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; RELEVANT KE:
US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT
                                                                                                                                                                                                                                                                                                     g
```

ö

ö

g à

```
AL INPORMATION.
APPLICANT: MAERTENS, GEERT
STUTYER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: NEW SEQUENCES AS PROPHYLACTIC, THERAPEUTIC AND DIAGN
                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                          ö
                                                        Score 135; DB 10; Length 103;
Pred. No. 8.1e-10;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 135; DB 10;
Pred. No. 8.5e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: 08/836,075
FILING DATE: CUNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 0ct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                            18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: KAMMERER, PATRICIA A.
                                                                                                                                                                                                                                                                                                    Sequence 14, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE: TEXAS
COUNTRY: USA
ZIP: 7710-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; Sequence 78, Application US/09921397
; Patent No. US20020151484A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STREET: P.O. BOX 4433
CITY: HOUSTON
                                                           86.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS
; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                           Query Match
Best Local Similarity 89.39
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 86.5
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                                           RESULT 7
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                     ò
                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       polypeptides selected from a
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      86.5%; Score 135; DB 10; Length 97; 89.3%; Pred. No. 7.6e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                 OFFINITY SIGNATURE SOLVE #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,875
                                                                          Sequence 8, Application US/09756875
Patent No. US20020150990A1
GENERAL INFORMATION:
APPLICANT: PIKE, IAN
TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 77, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
TITLE OF INVENTION: SID nucleic acids and pol
TITLE OF INVENTION: Pathogenic strain of the
TITLE OF INVENTION: applications thereof
FILE REFERENCE: B4809A - JA.
CURRENT FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: EP 00402225.7
PRIOR FILING DATE: 2000-08-03
NUMBER OF SECIENCE 1000-08-03
NUMBER OF SECIENCE 1000-08-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Mismatches
                                                                                                                                                                                                                                     ADDRESSEE: Suite 701-E Columbia Square STREET: 555 13th Street, N. W. CITY: Washington STATE: D. C. COUNTRY: U. S. ZIP: 20004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION:
CLASSIFICATION:
PRIOR APPLICATION DATA:
PREDICATION NUMBER: US 08/259,721
FILING DATE: 29-AUG-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTONNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 1808-157A
FREFREENCE/DOCKET NUMBER: 1808-157A
TELECHONE: (202)783-6040
INFORMATION FOR SEQ ID NO: 8:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 25; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-921-397-77
                                                             US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 77
```

```
AFFLICANT: FALLANCE, LIEVEN
STUTVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROFHYLACTIC, THERAPEUTIC AND DIAGNO
                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Microsoft Word 6.0 / ASCII text output
                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                          5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 135; DB 10;
Pred. No. 1.1e-09;
1; Mismatches 2;
                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                               NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                        5 PRPQRKTKRNTNRRPQDVKFPGGGOIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-899-046-152
; Sequence 152, Application US/09899046
; Publication No. US20030008274A1
                                                                                                                                                                                                           ; Sequence 60, Application US/09851138
; Publication No. US20020183508A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                         APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO: 60
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      86.5%;
                   25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 89.3°
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                 RESULT 10
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-851-138-60
                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 11
                                                             ò
                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STUTVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                     ö
             polypeptides selected from a
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTE: 1BM PC compatible
COMPUTE: 1BM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKnown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                       Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 137
                                                                                                                                                                                                                                                                                                                                                Score 135; DB 10; Length 1
Pred. No. 8.9e-10;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 135; DB 10;
Pred. No. 1.1e-09;
; TITLE OF INVENTION: SID nucleic acids and pol
TITLE OF INVENTION: pathogenic strain of the
TITLE OF INVENTION: applications thereof
FILE REFERENCE: B4809A - JAZ
CURRENT FILICATION NUMBER: US/09/921,397
CURRENT FILICATION NUMBER: EP 00402225.7
PRIOR APPLICATION NUMBER: EP 00402225.7
PRIOR FILING DATE: 2000-08 03 ...
NUMBER OF SEQ ID NOS: 156
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433 CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 PKPQKKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 46, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO: 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               86.5%;
89.3%;
                                                                                                                                                                                                                                                                           GRGANISM: Hepatitis C virus US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY: USA
ZIP: 77210-4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-851-138-46
                                                                                                                                                                                                                                                              TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 9
```

```
Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      inear
                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                  TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 14
US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ó
                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy. NUMBER OF SEQUENCES: 270 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 166;
                                                                                                                                                                                                                                                                                                                              Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
 COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                          Score 135; DB 11;
Pred. No. 1.3e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 135; DB 11;
Pred. No. 1.3e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred, No. 1.3e
1, Mismatches
                                                                                                                                                                                                                                                                                                                                                                   1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/09/878, 28: FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-878-281-152

Sequence 152, Application US/09878281

Publication No. US20030032005A1

GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 42, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                 PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                       FILING DATE:
INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                              86.5%;
89.3%;
                                                                                                                                                                                                          LENGTH: 166 amino acids
TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                              Query Match 86.5
Best Local Similarity 89.3
Watches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  , MOLECULE TYPE: protein US-09-878-281-152
                                                                                                                                                                                                                                                                  ; MOLECULE TYPE: protein US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
```

```
New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270\,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
COMPOTER: IBM PC Compatible
COMPOTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS.
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILIMG DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; DB 11; ; 1.3e-09; ches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                86.5%; Score 135; DB 11;
89.3%; Pred. No. 1.3e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   86.5%; Score 135; DE
89.3%; Pred. No. 1.36
iive 1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPOROPKRNTPRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 44, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-878-281-42
; Sequence 42, Application US/09878281
; Publication No. US20030032005A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: New sequence TITLE OF INVENTION: GENOTYPES FOR UNMBER OF SEQUENCES: 270 COMPUTER READABLE FORM: REDING TYPE: Floppy disk COMPUTER: IBM PC COMPUTED: COMPUTER: COMPUTER
                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERIFICS:
LENGTH: 169 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       169 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 89.32
Best Local Similarity 89.32
Best Local Similarity 89.32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           , MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT:
TITLE OF INVENTION: N
TITLE OF INVENTION: G
NUMBER OF SEQUENCES:
```

S G

Search completed: August 7, 2003, 12:01:13 Job time : 14.3636 secs

```
CNI4 fragment of H
HCV capsid peptide
HCV capsid peptide
HCV capsid peptide
Non-A, non-B hepat
Hepatitis C virus
HCO core-envelope
HCV cor
                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
Antigen pHCal01.
Antigen pHCbl01.
HCV fragment 1 / I
HCV fragment 2 / I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Non-A, Non-B Hepat
Non-A, Non-B Hepat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Antigenic epitope
HCV core protein 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Branched peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR29160
AAR29161
AAR41422
AAR41423
AAR51389
                                                                                                                                                                          AAY94410
AAY94409
                                                                                                                                                                                                                           AAR29534
AAR29535
                                                                                                                                                                                                                                                                                                                                               AAR69545
AAY15425
                                                                                                                                                                                                                                                                                                                                                                                  AAW01865
AAR12597
                                                                                                                                                                                                            AAR74563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY06675 standard; Protein; 28 AA
   AAY06675;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAYOG675
ID AAYO
   August 7, 2003, 11:05:37; Search time 38.5455 Seconds (without alignments) 115.301 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
/SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
/SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1981.DAT:*
/SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1981.DAT:*
/SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1985.DAT:*
/SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1986.DAT:*
/SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1986.DAT:*
/SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1986.DAT:*
/SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1980.DAT:*
/SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1980.DAT:*
/SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1980.DAT:*
/SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1980.DAT:*
/SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1990.DAT:*
/SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1990.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .DAT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.
                   GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                              al number of hits satisfying chosen parameters:
                                                                                                                                                                                                         154
1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                             1107863 seqs, 158726573 residues
                                                                                        protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Listing first 45 summaries
                                                                                                                                                                                                                                                                             Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                                                                                                                                                                                  Post-processing: Minimum Match 0% Maximum Match 100%
                                                                                                                                                                                                                                                                                                                                                                                  Minimum DB seq length: 0 · Maximum DB seq length: 20000000000
                                                                                                                                                                                          US-09-491-146A-25
```

BLOSUM62

Scoring table:

Searched:

Database :

Perfect score:

Title:

Sequence:

•

OM protein

Run on:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\* /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\* /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\* /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:\*

/SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:

/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998. /SIDS1/gcgdata/geneseg/genesegp-embl/AA1999 /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997

## SUMMARIES

Description	NC mosaic protein	Blood transmiscibl	HCV-S1 full-length	NC mosaic protein	HCV type 3 capsid	Hepatitis C virus	Hepatitis C virus	Hepatitis C virus	Hepatitis C virus
ΙD	AAY06675	AAR53417	AAE20477	AAY06673	AAB71258	AAR92968	AAR92969	AAR92971	AAR96547
DB	20	15	23	20	23	17	17	11	11
Length	28	3010	3010	28	189	191	191	191	319
% Query re Match Length DB II	100.0	100.0	100.0	94.8	94.8	94.8	94.8	94.8	94.8
Score	154	154	154	146	146	146	146	146	146
Result No.	1	7	m	4	2	9	7	æ	σ

Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination. New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual NC mosaic protein amino acid fragment C. (USSH ) US DEPT HEALTH & HUMAN SERVICES. 98WO-US17385 97US-0921887 17-JUN-1999 (first entry) Khudyakov YE; WPI; 1999-204671/17. Hepatitis C virus 25-AUG-1997; 21-AUG-1998; WO9910506-A1 04-MAR-1999. Fields HA,

```
ó
                                         homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosatic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for deceting a hepatitis infection and institution against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoassays. This provides improved detection of hepatitis C virus. Sequences ANYOGGS-633 represent amino acid sequence of each monomer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polymerase chain reaction; PCR; amplify; primer; non-A, non-B hepatitis; NANBH; virus; blood transmissible; detection; hepatitis virus; RT-PCR;
                                  ō
                                                                                                                                                                                                                                                                        Gaps
                                The invention relates to a mosaic protein, comprising a plurality
                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                             Length 28;
                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                             100.0%; Score 154; DB 20; 100.0%; Pred. No. 1.2e-14;
                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                     Mismatches
                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                           Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Len
                                                                                                                                                                                                                                                                                                                                                                                 AAR53417 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Blood transmiscible NANBHV protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C100 antibody; HCV RNA; NS5 region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      <u>r</u>en
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  'label- His, Arg
                                                                                                                                                                                                                                                                     ö
       Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  non-B hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label- Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 label- Ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /label- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /label= Phe,
Misc-difference 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 'label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label Thr,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  label- Gln,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           label- Lys,
                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                           comprising the NC mosaic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 438
                                                                                                                                                                                                                                                       st Local Similarity atches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 311
                                                                                                                                                                                                                    28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                17-JAN-1995
                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                             very Match
                                                                                                                                                                                                                                                                                                                                                                                                          AAR53417;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Non-A,
                                                                                                                                                                                                                                                                                                                                                        RESULT 2
AAR53417
ID AAR5
```

THE SOUND STATE OF THE STATE OF

8

```
This sequence is encoded by the genome of a blood transmissible non-A, non-B hepatitis (NANBH) virus. The CDNA sequence was isolated using the primers given in AAQG3500-35. The amplified fragments are used in the detection of hepatitis virus. The target DNA was isolated from serum of chronically infected NANBH patients who were C100 antibody-positive and HCV RNA (NS5 region) positive. Reverse transcription-PCR and PCR were performed on CDNA and the total human NANBH DNA was constructed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Blood-transmissible non-A non-B hepatitis virus DNA - used detection of hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Page 8-20; 22pp; Japanese.
                                                                                                                                                                                                                                                                                                                    Asp
                                                                                                                                                                                                                                                                                                                                                           Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                    'label- Gly, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Met
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Met
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Leu
                        /label= Leu,
Misc-difference 1017
                                                                                                                                                         label- Glu,
                                                                                                                                                                                                                                     'label- Met,
                                                                                                                                                                                                                                                                                                                /label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          'label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /label Tyr,
/label- Arg,
                                                                                                                 /label- Thr,
                                                                                                                                                                                                                                                                             label- Asn,
                                                                                                                                                                                                                                                                                                                                                       label- Asn,
                                                                                                                                                                                                                                                                                                                                                                                               'label- Ala,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'label- Glu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label - Val,
                                                                                                                                                                                              label- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            'label- Lys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           label - Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  'label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /label- Ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92JP-0051885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92JP-0051885
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 1896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 2844
                                                                                                                                                                                                               Misc-difference 1205
                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 1895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 2486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 2488
                                                                                                                                    Misc-difference 1056
                                                                                                                                                                           Misc-difference 1201
                                                                                                                                                                                                                                                                                                                                     Misc-difference 1455
                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 1828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 2143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 2462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 2968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 2989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 2990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 2144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1994-163130/20.
N-PSDB; AAQ63499.
                                                                                                                                                                                                                                                                                           Misc-difference 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 286
                                                                                             Misc-difference 103
                Misc-difference 759
                                                                                                                                                                                                                                                      disc-difference 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (KAEN/) KAENNO K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JP06105690-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 1;
```

Sequence Query Match

SXS

```
The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vacchation against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 mosaic protein, comprising a plurality of homologous antigenic tides from different genotypes of a species - useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 146; DB 20;
Pred. No. 1.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 peptides from different genotypes of a species detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28
                                   28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NC mosaic protein amino acid fragment A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPORKTKRNTYRRPODVKFPGGGQIVG
                                                                           PKPQRKTKRNTYRRPQDVKFPGGGQIVG
                                   PKPQRKTKRNTYRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB71258 standard; protein; 189 AA.
                                                                                                                                                                                                                                                                                                               Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                       AAY06673 standard; Protein; 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 5; Fig 9; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94.8%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98WO-US17385.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                           17-JUN-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity 96.4 les 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1999-204671/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO9910506-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              04-MAR-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                   AAY06673;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB71258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Fields
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 5
AAB71258
                                                                                                                                                                                                                                                               AAY06673
                                                                                                                                                                                                                                                                                                                                                                                   NAME OF THE OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                   οχ
                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   expression cassette including a first polynucleotide region including a 5' non-coding region (NCR) sequence of an RNA virus and at least an N-terminal portion of a coding sequence of RNA virus, a second polynucleotide region including a 3' untranslated region (UTR) sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus and at least a C-terminal portion of a coding sequence of the virus and at least a C-terminal portion of a coding sequence of the virus and at least a C-terminal portion of a coding sequence of the virus and at least a C-terminal portion of a coding sequence of the virus and a third polynucleotide regions; and a promoter sequence being operatively linked to expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette. Nucleic acid construct of the invention is useful for detecting the presence of an RNA virus in a cell. It is also useful for screening anti-viral drugs and determining drug in the construct of the invention and construct of the invention o
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Novel nucleic acid construct useful for detecting the presence of RNA virus, comprises an expression cassette and a promoter operably linked to expression cassette for minus strand RNA transcription of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          an RNA virus. The present sequence is Hepatitis C virus HCV-S1 full-length polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nucleic acid construct; expression cassette; non-coding region; NCR; untranslated region; UTR; anti-viral drug; drug resistance; HCV-S1; Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                               Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                               Indels
                                                                                                                                                   ; DB 15;
1.4e-12;
hes 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0%; Score 154; DB 23;
100.0%; Pred. No. 1.4e-12;
ive 0; Mismatches 0;
                                                                                                                                               / Match 100.0%; Score 154; C
Local Similarity 100.0%; Pred. No. 1.4
les 28; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                   INST MOLECULAR & CELL BIOLOGY. EHRLICH G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Example 1; Page 70-81; 81pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAE20477 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hong WJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV-S1 full-length polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20-JUL-2001; 2001WO-IL00669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24-JUL-2000; 2000US-220248P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Tan YH, Lim SP, Lim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (HCV) isolate HCV-S1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2002-280605/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3010 AA;
                                                                           3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAD33038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200208447-A2.
from 23 clones.
```

(MOLE-) INST (EHRL/) EHRL)

31-JAN-2002.

01-JUL-2002

AAE20477;

RESULT 3 AAE20477

ò qq ö

Gaps

ö

resistance of

Sequence

Length 28; Indels

```
El; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA and amino acid sequence of HCV envelope 1 and core proteins -used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR92936-R92987 are HCV core proteins derived from 52 different HCV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Match 94.8%; Score 146; DB 17; Length 191; Local Similarity 96.4%; Pred. No. 1.1e-12; les 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus isolate S52 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
(USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR92969 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 4; Page 207; 340pp; English,
                                                                                                                                                                                                                                                                                                                                                                       Bukh J, Miller RH, Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Purcell
                                                                                                                                                                                         95WO-US10398.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1996-139709/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1996-139709/14
                                Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAT16642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO9605315-A2.
                                                                                                                                                                                                                                          15-AUG-1994;
                                                                                                                                                                                         15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-AUG-1995;
                                                                                                                                        22-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     nepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR92969:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR92969
        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This interior describes a mover accounted that it is a conservity described and described that the consecutive amino acids from the capsid protein, provided that the consecutive amino acids from the capsid protein. The action of the invention has virucide, antidiammatory and hepatotropic activity. The attendated vaccine, and similar nucleic acid vaccines that encode the mutated capsid protein, are useful for protection against a choice for flavivitus diseases, e.g. yellow fever, Japanese and hepatitis, dengue, classical swine fever, bovine viral diarrhoea and hepatitis C. The specified deletion: (i) produces a reliably attendated contracted and does not revert to virusence; (ii) is exactly defined and does not effect immune responses to important proteins; and (iii) can not generate a non-natural virus by recombination. The mutant viruses eliminate the need to produced and the sexpense. The protective response to flavivirus lasts significantly longer than that to killed vaccines.

This sequence represents a fragment of the capsid protein from Hepatitis cribs (ivus) (ivus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Attenuated flavivirus live vaccine, useful for protection against e.g. yellow fever, comprises virus with attenuating deletion of amino acids from the capsid protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This invention describes a novel attenuated flavivirus live vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                       Capsid protein; attenuated vaccine; virucide; antinflammatory;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 189;
                                                                                                                                  hepatotropic; yellow fever; Japanese encephalitis; dengue; classical swine fever; bovine viral diarrhoea; hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              94.8%; Score 146; DB 23;
96.4%; Pred. No. 1.1e-12;
ive 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus isolate HK10 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV type 3 capsid protein fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR92968 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Disclosure; Fig 2; 30pp; German.
                                                                                                                                                                                                                                                                                                                                                                  11-FEB-2002; 2002WO-AT00046.
                                                                                                                                                                                                                                                                                                                                                                                                                   21-FEB-2001; 2001AT-0000272.
  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match
Best Local Similarity 96.45
Matches 27: Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           02-OCT-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2002-667064/71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mandl C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (HEIN/) HEINZ F X. (MAND/) MANDL C.
                                                                                                                                                                                                            Hepatitis c virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               189 AA;
                                                                                                                                                                                                                                                             WO200266621-A1.
18-NOV-2002
                                                                                                                                                                                                                                                                                                              29-AUG-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Heinz FX,
```

Sequence

hepatitis.

AAR92968;

RESULT 6 AAR92968

ò 용

ö

Gaps

ô

ò g

```
/label= Gly
/note= "amino acid in this position is designated X in
the specification, but codon usage shows that the
only possible amino acid at this pos. is Gly*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "amino acid in this position is designated X in
the specification, but codon usage shows that the
only possible amino acid at this pos. is Gly"
                                                                                                                                                                                                                                                                                                                Hepatitis C virus; subtype; polymerase chain reaction; amplification; PCR; primer; probe; antibody; infection.
                                              Gaps
                                           ;
0
                                                                                                                                                                                                                                                                                    Hepatitis C virus types 7c(8a) isolates VN4 amino acids 1-317.
           Length 191;
                                            Indels
           Score 146; DB 17;
Pred. No. 1.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  .44..149
'label" Val, Ala, Glu, Gly
                                              Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              G1y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       'label- Val, Ala, Glu, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      'label- Cys, Arg, Ser, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Val
                                                                            1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                            Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Lys,
                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers Misc-difference 144
                                                                                                                                                                                       Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /label- Met, Thr,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              'label- Val, Ala,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Leu
                                                                                                                                                                                                                                                                                                                                                                                                                               label- Met, Leu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          'label- Ser, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /label- Asn, Asp
                                            ö
                                                                                                                                                                                       AAR96547 standard; peptide; 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /label- Met,
           94.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          95EP-0870076.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95WO-EP04155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /label= Gly
                                                                                                                                                                                                                                                     (first entry)
        Query Match
Best Local Similarity 96.4°
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (INNO-) INNOGENETICS NV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Stuyver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1996-251460/25.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference 14
                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO9613590-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23-OCT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUN-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-OCT-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Maertens G,
                                                                                                                                                                                                                                                     10-MAR-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09-MAY-1996
                                                                                                                                                                                                                       AAR96547;
                                                                                                                                                        RESULT 9
                                                                                                                                                                                                      Q
                                                                             ö
                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, sallva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                      AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                envelope 1 and core proteins - as vaccines against HCV infection
                             envelope 1 and core proteins - as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                            Score 146; DB 17; Length 191;
Pred. No. 1.1e-12;
0; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus isolate DK12 core protein.
                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (USSH ) US DEPT HEALTH & HUMAN SERVICES.
(USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 4; Page 209-210; 340pp; English.
                             amino acid sequence of HCV determine HCV genotype and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid sequence of HCV determine HCV genotype and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR92971 standard; Protein; 191 AA
                                                                          Claim 4; Page 208; 340pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bukh J, Miller RH, Purcell RH;
                                                                                                                                                                                                                                                                                   Query Match 94.8%;
9est Local Similarity 96.4%;
atches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1996-139709/14.
                                                                                                                                                                                                                                                      191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAT16645
N-PSDB; AAT16643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                02-0CT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR92971;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 used to
                               DNA and
                                                nseq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR9297
```

Gaps

ö

Indels

Length 36;

38888

a à

EXXEXEX SXI

```
The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 3B and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contained in the capsid protein of the virus and themselves contained in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              94.2%; Score 145; DB 14; Length 38; 96.4%; Pred. No. 3e-13; .ive 0; Mismatches 1; Indels
        HCV infection and to raise antibodies against it.
                                                                                          Score 145; DB 16;
Pred. No. 2.9e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                    1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                           1 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                           AAR30687 standard; peptide; 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                      94.2%;
ilarity 96.4%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91US-0714471.
91US-0718052.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity 96.4%;
les 27; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92WO-US03635
                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV capsid peptide No. 23.
                                                                                                                                                                                                                                                                                                                                                                                               (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Jolley ME, Leahy DC,
                                                                                Query Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1993-018073/02.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO9222571-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                         25-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                  11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           antibodies.
                                                                                                                                                                                                                                                                                                                                                     AAR30687;
                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                         RESULT 11
          SXS
                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                               The sequences AAR86526-R9657B represent novel sequences isolated from hepatitis C virus subtypes different from subtypes la-c, 2a-d, 3a-f, 4a-f, 5a and 6a. Theye sep. from the novel subtypes id-f, 2e-i, 2k, 2l, 3d, 4k-m, 7a-c or types 9. from the novel subtypes id-f, 2e-i, 2k, 2l, 3d, 4k-m, 7a-c or types 9. from the novel subtypes id-f, 2e-i, 2k, 2l, 3d, 4k-m, 7a-c or types 9. from the Corresp. to the 5' cuntranslated region (UR), the Core/El, NS4 or NSB regions of the St sequence represents amino acids 1-317 from the HCV types were isolated from patients with chronic HCV from the Benelux countries, France, Cameroon and Vietnam, because of their absentint reactivities. The RNA was extracted, cDNA synthesised and PCR amplified, cloned and genotyped. The 5'UR, Core/El and NSSB regions were viruses into (subtypes based on comparison with known sequences. The sequences can also be used to synthesise probes and primars for the sequences can also be used to synthesise probes and primars for the sequences can also be used to synthesise probes and primars for the canti-HCV antibodies, for HCV typing or to prevent HCV infections.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This sequence is encoded by the oligonucleotide, CN14, and represents the peptide fragment CP14. CP14 is a fragment of the core region of hepatitis C virus (HCV). CP14 may be used in the detection of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                 Hepatitis C virus poly:nucleic acid unique to unidentified sub:type - used to develop probes and primers for new sub:types and vaccines to prevent and treat infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Detection of hepatitis C virus - using oligopeptide fragment of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CN14; CP14; core region; hepatitis C virus; HCV; detection;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CN14 fragment of HCV core region.
                                                                                                                   Claim 25; Fig 3; 150pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAW06487 standard; Protein; 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 1; Page 6; 7pp; Japanese
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93JP-0156026.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      93JP-0156026.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  31-JAN-1997 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (IMMO ) IMMUNO JAPAN KK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1995-047903/07.
N-PSDB; AAT45055.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            319 AA;
N-PSDB; AAT27958.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JP06327482-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29-NOV-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            core region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW06487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          antibody.
```

ö

Gaps

ö

1;

AAR30688;

```
The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptities are contained in the capsid protein of the virus and themselves contain epitope artibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                                         Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              non-B hepatitis virus; NANBHV; hepatitis C virus; HCV; SNV; NS1; NS2; NS3; antigen; detection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 145; DB 14; Length 38;
Pred. No. 3e-13;
0; Mismatches 1; Indels
non-A, non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Non-A, non-B hepatitis virus corel region fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (SANW ) SANWA KAGAKU KENKYUSHO CO.
(TOFU ) TONEN CORP.
(TOKR-) 2H TOKYOTO RINSHO IGAKU SOGO KENKYUSHO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPORKTKRNTNRRPODVKFPGGGOIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                  Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR54065 standard; Protein; 38 AA
                                                                                                                                                                                                                                                                      Todd JA;
                                                                                                                                                                                                                                   (BAXT ) BAXTER DIAGNOSTICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 94.2%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                             91US-0714471.
91US-0718052.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92JP-0088140,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  92JP-0088140.
                                                                                                                                           92WO-0S03635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           14-FEB-1995 (flrst entry)
                                                                                                                                                                                                                                                                    Leahy DC,
                                                                                                                                                                                                                                                                                                      WPI; 1993-018073/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12-MAR-1992;
                                                                      W09222571-A1
                                                                                                                                           29-APR-1992;
                                                                                                                                                                             13-JUN-1991;
20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JP06141870-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24-MAY-1994.
                                                                                                          23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     core; ENV;
                                                                                                                                                                                                                                                                      Jolley ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR54065;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR54065
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope groups. These peptides can be used in immnosassays for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH;
                                                                                                                                                                                            Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          94.2%; Score 145; DB 14; Length 38; 96.4%; Pred. No. 3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPORKTKRNTYRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Flg 1F; 66pp; English.
                                       AAR30688 standard; peptide; 38 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR30689 standard; peptide; 38 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                                                              92WO-US03635
                                                                                                                                                                                                                                                                                                                                                                                                                  91US-0714471
91US-0718052
                                                                                                          (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                              HCV capsid peptide No. 24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 96.4'
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV capsid peptide No. 25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Jolley ME, Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1993-018073/02
                                                                                                                                                                                                                                                                      Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38 AA;
                                                                                                                                                                                                                                                                                                                                                                              29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                  13-JUN-1991;
20-JUN-1991;
                                                                                                        25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                        WO9222571-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                           23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  antibodies
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
```

AAR30689;

8

ò

ö

Gaps

Search completed: August 7, 2003, 11:14:06 Job time : 38.6364 secs

us-09-491-146a-25.rag

```
;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present sequence is that of an antigen used in a novel assay for determining anti-HCV antibody titre. A sample is diluted to an absorbance of not more than 2.0 and the peptide antigen is added. The method is useful for monitoring the effect of interferon treatment on anti-HCV antibody titres.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                       DNA coding a Non-A, non-B hepatitis virus antigen - useful for detecting HCV within serum
                                                                                                                                                                                                                                           ö
                                                                                                                        Hepatitis C virus #4 and #6 genes were isolated (AAQ64068-69). Both genes contain the core, ENV, NS1, NS2 and NS3 regions. A core region fragment is given in AAQ64067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Determination of the antibody titre against hepatitis C virus antigen - using dilute solution and HCV antigenic peptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV; non-A, non-B hepatitis virus; antigen; immunoassay;
interferon treatment; monitoring; antibody titre; detection.
                                                                                                                                                                                                        Query Match 94.2%; Score 145; DB 15; Length 38; Best Local Similarity 96.4%; Pred. No. 3e-13; Matches 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 94.2%; Score 145; DB 16; Length 40; Best Local Similarity 96.4%; Pred. No. 3.2e-13; Matches 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus core antigenic peptide.
                                                                                                                                                                                                                                                                1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Page 11; 22pp; Japanese,
                                                                                                                                                                                                                                                                                                                                                              AAR84559 standard; peptide; 40 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 1; Page 10; 11pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      94JP-0082160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 94JP-0082160
                                                                                                                                                                                                                                                                                                                                                                                                                    06-JUN-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (ARIM/) ARIMA T.
(TORA ) TORAY IND INC.
             WPI; 1994-205028/25.
N-PSDB; AA064067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1995-386063/50.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
                                                                                                                                                                                38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    40 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JP07260792-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-OCT-1995.
                                                                                                                                                                                 Seguence
                                                                                                                                                                                                                                                                                                                                                                                             AAR84559;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                    RESULT 15
AAR84559
ઠ
                                                                                                                                                                                                                                                                                                                                                                 ô
```

a

```
August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                 .al number of hits satisfying chosen parameters:
                                                                                                                                                                                       US-09-491-146A-25
154
1 PKPQRKIKRNIYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                            283308 segs, 96168682 residues
                                                                          OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                    BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                         Title:
Perfect score: ]
Sequence:
                                                                                                                                                                                                                                                                    Scoring table:
                                                                                                                                                                                                                                                                                                                              Searched:
                                                                                                              Run on:
```

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries Minimum DB seq length: 0
Maximum DB seq length: 2000000000 PIR\_76:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description	1011411000	qenome polyprotein																				genome polyprotein		genome polyprotein	genome polyprotein	genome polyprotein				o
	10		A45573	S41359	S41358	S41360	PC2061	S41288	S21336	S41353	S41355	S41357	S41356	S41348	S41371	S41341	S41370	S41369	S41368	S41342	S41344	S41350	S41354	S41345	S41347	S41343	S41346	S21471	S12707	PC1284	JQ1925
	DB	:	-	~	~	~	7	7	~	~	~	7	7	7	7	7	~	7	7	~	7	~	~	7	~	~	~	~	~	~	7
	Length		3010	114	114	124	7	492	88	108	108	108	108	108	112	112	114	114	114	115	115	115	115	115	115	115	118	369	441	513	~
ď	Query		100.0	4.	4.	94.8	94.8	4.	94.2	94.2	94.2		94.2	٠	94.2		٠	94.2	•	4	94.2	94.2	94.2	4	94.2	94.2		÷	94.2	94.2	94.2
	Score		154	146	146	146	146	146	145	145	145	145	145	145	145	145	145	145	145	145	145	145	145	145	145	145	145	145	145	145	145
	Result No.		г	7	m	4	Ŋ	9	7	80	6	10	. 11	12	13	14	15	. 16	17	18	19	20	21	22	23	24	25	56	27	28	29

polyprotein - hepa genome polyprotein	•
JQ1926 JH0711 S18031 S18031 S18031 S18032 PN0677 PO0883 JQ0883 JQ0883 S18030 S18030 S40770 JC5563 JQ1303	
00000000000	
523 550 782 782 782 782 782 874 874 876 3010 3010 3010 3011 3011	
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%	
145 145 145 145 145 145 145 145 145 145	
0 H B B B B B B B B B B B B B B B B B B	

## ALIGNMENTS

283308

RESULT 2 S41359 genome polyprotein - hepatitis C virus (genotype 3, N2) (fragment) N;Contains: core protein C;Species: hepatitis C virus A;Variety: genotype 3, N2 C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000 C;Accession: S41359

```
ö
                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A)Cross-references: EMBL:229461; NID:9443884; PIDN:CAA82599.1; PID:9443885
A)Experimental source: genotype 3, N1
C;Superfamily: hepartitis C virus genome polyprotein
C;Seywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: genomic RNA
A; Mostaces: 1-124 CVAN>
A; Residues: 1-124 CVAN>
A; Cross-references: EMBL: 229463; NID: 9443888; PIDN: CAA82601.1; PID: 9443889
A; Experimental source: genotype 3, N3
A; Experimental source: genotype 3, N3
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-124/Product: core protein #status predicted <MAT>
                                                                                                                                                               A)Cross-references: EMBL:229462; NID:9443886; PIDN:CAA82600.1; PID:9443887
A)Experimental source: genotype 3, N2
A)Experimental source: genotype 3, N2
C)Superfamily: hepatitis C virus genome polyprotein
C)Superfamily: hepatitis C virus genome polyprotein
C)Sewyords: capsid protein; core protein; polyprotein
F):1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                     by LiPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Lipa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lipa
                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus (genotype 3, N3) (fragment)
N;Contains: core protein
N;Contains: core protein
S;Species: hepatitis C virus
A;Variety: genotype 3, N3
C;Date: 19-May-1994 *sequence_revision 26-Jul-1996 *text_change 17-C;Date: 19-May-1994 *sequence_revision 26-Jul-1996 *text_change 17-C;Date: 19-May-1994 *sequence_revision 26-Jul-1996 *text_change 17-C;Date: 19-May-1994 *sequence_revision 26-Jul-1994
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Ancestion: S41360
A; Accession: S41360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatitis C virus (genotype 3, N1) (fragment) N:Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ργ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               #text_change
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 114;
                                                                                                                                                                                                                                                                                                                                          Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        to 5
     Library, January 1994
hepatitis C virus genotypes 1 to 5
                                                                                                                                                                                                                                                                                                                                                                                              1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A Variety: genotype 3, NI
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_(C;Accession: S41358
C;Accession, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1
                                                                                                                                                                                                                                                                                                                                    Score 146; DB 2;
Pred. No. 1.5e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 146; DB 2;
Pred. No. 1.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          94.8%;
96.4%;
                                                                                                                                                                                                                                                                                                                                    Query Match 94.8%;
Best Local Similarity 96.4%;
Matches 27; Conservative
submitted to the EMBL Data Lib
A;Description: Analysis of hep
A;Reference number: $41341
A;Accession: $41359
A;Molecule type: genomic RNA
A;Residues: 1-114 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A Accession: S41358
A Molecule type: genomic RNA
A Residues: 1-114 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ٤
```

Length 124;

DB 2;

Score 146;

Query Match

```
Grontains: envelope protein E1; nonstructural protein E2/NS1
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Species: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Nov-2000
C;Accession: PC2061
R;Li, J.S.; Vitvieski, L.; Tong, S.P.; Trepo, C.
Blochem. Biophys. Res. Commun. 199, 1474-1481, 1994
A;Title: Identification of the third major genotype of hepatitis C virus in France. A;Reference number: PC2060; MUID:94197744; PMID:8147893
A;Accession: PC2061
A;Molecule type: mRNA
A;Residues: 1-411 <LLIJ>
A;Residues: 1-411 <LLIJ>
A;Residues: 1-411 <LLIJ>
A;Cross references: GB:L12355; NID:9410169; PIDN:AAA20155.1; PID:9410170
C;Superfamily: hepatitis C virus genome polyprotein
C;Reynords: ATP; capsid protein; envelope protein E1*status predicted <NPE>
F;384-411/Product: envelope protein E2/NS1 #status predicted
F;384-411/Product: nonstructural protein E2/NS1 #status predicted
F;196,209,234,305,325/Binding site: carbohydrate (Asn) (covalent) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nonstructural protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sylicon
Sylico
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome polyprotein S29 (core protein region) - hepatitis C virus (fragment) C;Species: hepatitis C virus C;Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 17-Nov-2000 C;Accession: S21336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 492;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 411;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Indels
                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Indels
. 1.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 146; DB 2;
Pred. No. 6.3e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Match 94.8%; Score 146; DB 2; Local Similarity 96.4%; Pred. No. 5.3e-13; les 27; Conservative 0; Mismatches 1;
                                       Pred. No. 1.6e
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    32
                                                                                                                                                                                                                                          32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                               PKPQRKTKRNT IRRPODVKFPGGGOIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                           polyprotein N2 - hepatitis C
                                  96.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94.8%;
ilarity 96.4%;
Conservative
                                       Best Local Similarity 96.49
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
nes 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              $21336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S41288
```

```
N; Contains: core protein
C; Species: hepatitis C virus (genotype 2, N5) (fragment)
C; Species: hepatitis C virus
A; Variety: genotype 2, N5
A; Variety: genotype 2, N5
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 84136
B; van Doorn, L.J.; Rleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                                                                                                                                                                                 A; Variety: genotype 2, N6
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 541357
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (genotype 2, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Experimental source: genotype 2, N5
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-108/Product: core protein *status predicted <NAT>
             1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    c.neyworus: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 145; DB 2;
Pred. No. 1.9e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 145; DB 2;
Pred. No. 1.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A:Experimental source: genotype 2, N6
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotei
             Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Mismatches
                                                                                       5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PKPQRKTKRNTCRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                            1 PKPORKTKRNTYRRPQDVKFPGGGGIVG
             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 94.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      96.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Molecule type: genomic RNA
A;Residues: 1-108 <VAN>
A;Cross-references: EMBL:229460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
A; Cross-references: EMBL: 229459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 96.4 Matches 27; Conservative
             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Reference number: S41341
                                                                                                                                                                                                                                                                   N;Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Accession: S41356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Accession: S41357
       27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S41348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                            ò
                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
NyContrains: core protein
C:Species: hepatitis C virus
Ayariety; gencype 2, Na
C:Species: hepatitis C virus
C:Species: hepatitis C virus
C:Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C:Accession: $41355
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A; Reference number: $41341
A; Reference number: $41341
A; Reference number: $41341
A; Residues: 1-108 <VAN>A; Cross-reference genonic RNA
A; Residues: caference genotype 2, N4
C:Superimental source: genotype 2, N4
C; Superimental source: genotype 2, N4
C; Superimental capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <WAT>
                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (genotype 2, N2) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N2
C;Species: 19-May-1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-200C
C;Date: 19-May-1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-200C
C;Submitted to the EMBL Data Library, January 1994
A;Reference number: S41341
A;Reference number: S41341
A;Residues: 1-108 </AN
A;Residues: 1-108 </AN
A;Cross references: EMBL:229456
A;Experimental source: genocype 2, N2
C;Superiamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; status predicted <AMT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94.2%; Score 145; DB 2; Length 108; llarity 96.4%; Pred. No. 1.9e-13; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 108;
                                                                                                                                                                                                                                                                                                                                           1; Indels
                                                                                                                                                                                                                                                                                      Score 145; DB 2;
Pred. No. 1.6e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 145; DB 2;
Pred. No. 1.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                   11 PKPQRKTKRNTSRRPQDVKFPGGGQIVG 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94.2%;
                                                                                                                                                                                                                                                                                      Query Match 94.2%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
nes 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ş
```

ö

F;1-112/Product: core protein #status predicted <MAT>

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000 C;Accession: S4134 B sequence_revision 26-Jul-1996 #text_change 17-Nov-2000 C;Accession: S4134 B submitted to the EMBL Data Library, January 1994 A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA. A;Reference number: S4134 A;Reference number: S4134 A;Reference number: S4134 A;Residues: 1-108 <-VAN> A;Residues: 1-108 <-VAN> A;Residues: 1-108 <-VAN> A;Cross references: EMBL: Z29451 A;Experimental source: genotype 1, N6 C;Superfamily: hepatitis C virus genome polyprotein C;Superfamily: hepatitis C virus genome polyprotein C;Reywords: capsid protein; core protein; polyprotein F;1-108/Product: core protein #status predicted <-Mat>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              %).Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N5
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (genotype 1, NI) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, NI
C;Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Rivan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
A;Accession: S41341
A;Accession: S41341
A;Molecule type: genomic RNA
A;Residues: 1-112 cVAN>
A;Cross-references: EMBL: Z2944; NID: 9443850; PIDN: CAA82582.1; PID: 9443851
A;Experimental source: genomypel, NI
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA
A; Reference number: S4134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 145; DB 2; Length 112;
Pred. No. 2e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A Accession: S41371
A Molecule type: genomic RNA
A; Readures: 1-12 C4NA>
A; Coss - references: EMBL: 229474
A; Experimental source: genotype 5, N5
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capaid protein, core protein; polyprotein
F; 1-112/Product: core protein #status predicted <NAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 145; DB 2;
Pred. No. 1.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 96.4 Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Accession: S41371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
```

```
ö
                                            ö
                                                                                                                                                                                                                                               N;Contains core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N4
C;Date: 19-May_1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                            A; Accession: S41370
A; Molecule type: genomic RNA
A; Residues: 1-114 <VAN>
A; Cross-references: EMBL: Z29473; NID: 9443908; PIDN: CAA82611.1; PID: 9443909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                            LIPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                            ö
                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                            βy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 114;
  Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                            Ç
                                                                                                                                                                                                                                                                                                                                             R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1
A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A.Experimental source: genotype 5, N4 (5) Substânily: hepatitis C virus genome polyprotein C.Keywords: capsid protein; core protein; polyprotein F;1-114/Product: core protein *status predicted 
    DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    94.2%; Score 145; DB 2; 96.4%; Pred. No. 2e-13; iive 0; Mismatches
                     2e-13;
  Score 145; DB
Pred. No. 2e-1
0; Mismatches
                                                                                                      1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7, 2003, 11:21:47
Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 96.4 nes 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Search completed: August
Job time: 9.54545 secs
                                                                                                                                                                                                                                                                                                                              C; Accession: 541370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loc
Matches
                                                                                                               . <u>අ</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec Run on:

US-09-491-146A-25

1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28 Perfect score: Seguence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 tal number of hits satisfying chosen parameters:

127863 seqs, 47026705 residues

Searched:

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries Minimum DB seq length: 0 Maximum DB seq length: 2000000000

SwissProt\_41:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Description		hepatitis	Q01404 hepatitis c	Q01403 hepatitis c	P27960 hepatitis c			ď	၎	4	ב	ď	h genome	homo sapi		pos	-		O09000 m nuclear r		P50894 fugu rubrip	homo s	herpes	-	P38261 saccharomyc	O31162 spiroplasma	Q8r0n9 mus musculu	23	P48155 manduca sex	22	37	P75278 mycoplasma	~
ID	POLG_HCVJT	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVTW	POLG_HCV1	FURI_HUMAN	FURI_MOUSE	FURI_BOVIN	SYN_BRUMA	FURI_RAT	NCO3_MOUSE	RK4_TOBAC	RS7_FUGRU	RS7_HUMAN	IE63_HSV2H	B3AT_CHICK	YBV2_YEAST	RL16_SPICI	ZDH1_MOUSE	YB1_XENLA	RS7_MANSE	RS7_XENLA	RK4_SPIOL	YF08_MYCPN	PDPK_HUMAN
DB	٦	٦	-	ч	7	7	П	-	~	-	-	H	-	Н	-	-	-	-	-	-	Н	-	-	H	-	-	-	Н	-	-		٦	~4
% Query Match Length	3010	513	520	520	737	737	3010	3010	3011	3033	3033	3010	3011	794	6	797	548	793	1398	282	194	194	512	a	753	137	485	303	190	σ	σ	209	2
% Query Match	0	4	4	4	◂	◂	4	94.2	4	4	4	9	σ	σ	S	Š.	ω.	m.	e,	ä	31.8	ä	ij	Ξ.	ö	ö	ö	ö	ę.	6	6	29.9	o.
Score	154	145	145	145	145	145	4	145	4	4	145	138	137	9	52	54		51	51	20	49	49		4	47.5	47	4	46.5	46	46	46	46	46
Result No.	-	7	e	₹	S	9	7	89	σ	10	11	12	13	14	15	16	17	18	19		21	22	23	24				28.					33

MEROPS; 229.001; -. MEROPS; U39.001; -. InterPro; IPR001410; DEAD.

Q9z2a0 mus musculu	O9z848 chlamydia p	P51816 homo sapien	P19711 bovine vira	Q9vrj8 drosophila	Q10101 schizosacch	P48431 homo sapien	P48432 mus musculu	P54231 ovis aries	P18447 murine coro	P03416 murine coro
PDPK_MOUSE	RNR CHLPN	FMR2_HUMAN	POLG_BVDVN	U139_DROME	RS7_SCHPO	SOX2_HUMAN	SOX2_MOUSE	SOX2_SHEEP	NCAP_CVM3	NCAP_CVMA5
	٠,		-	-	н	н	٦	-	-	٦
50 50 50 50 50 50 50 50 50 50 50 50 50 5	676	1311	3988	108	195	317	319	320	454	454
29.0	29.6	29.9	29.5	29.5	29.5	29.5	29.2	29.5	29.5	29.5
	9	46	45.5	45	4.5	45	45	45	45	4.5
4 4	_		4							

## ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
Makazawa T., Nato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Nakazawa T., Hijikata M., Ishimura Y., Shimotchno K.;
"Molecular cloning of hepatitis C virus genome from a single Japanese cartaler: sequence variation within the same individual and among infected individuals.";
Virus Res. 23:39-53(1992).
L. FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE WEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C -1 CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
C -1 CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                         01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last annotation update)
66-2003 (Rel. 42, Last annotation update)
66-2004 (Rel. 42, Last annotation update)
67-2005 (Rel. 42, Last annotation update)
67-2006 (MSI); Protein P7; Nonstructural protein NS2 (PS2);
67-2006 (MSI); Protease/helicase NS3 (PP0) (Hepacivirin)
65-3-4-21-3; Protease/helicase NS3 (PP0) (Hepacivirin)
65-3-4-21-39; Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P6); Nonstructural protein NS5A (P6); Nonstructural protein NS5A (P6) (PP0) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
69-60 (PP0) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
69-61 (PP0) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (RNA)(N).
SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROFEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E, THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
       PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; D11168; BAA01943.1; -.
       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR; A45573; A45573.
PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=31642;
POLG_HCVJT
Q00269;
                                                                                                                                                                                                                                                                                                                                             Hepacivirus
```

01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
60-OCT-2001 (Rel. 40, Last annotation update)
Envelope 91ycoprotein El (GP32) (GP35); Envelope 91ycoprotein El (GP32) (GP35); Envelope 91ycoprotein El (Hepatitis C virus (1solate HC-J2) (HCV).
Hepatitis C virus (1solate HC-J2) (HCV).
Hepatitis C virus (1solate HC-J2) (HCV).

SEQUENCE FROM N.A.

BEDLINE-92330232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
Tranaka T., Fukuda S., Tsuda F., Mishiro S.;
Trill-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes.";

NCBI\_TaxID=11111;

'n

STANDARD;

POLG\_HCVJ2 P27959;

POLG\_HCVJ2

RESULT 2

```
REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/EZ (POTENTIAL).
NON STRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RAM-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100.0%; Score 154; DB 1; Length 3010; 100.0%; Pred. No. 1.3e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       326573 MW; 94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LINKED (GLCNAC.
LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                 InterPro: IPR002868; HCV_NSSA.
InterPro: IPR001065; HCV_NSSA.
InterPro: IPR001095; RNA_POL_DS_PS.
InterPro: IPR001094; RNA_POL_DS_PS.
InterPro: IPR001094; RNA_POL_PSVIr.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01559; HCV_env; 1.
Pfam; PF01559; HCV_NS1; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00271; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DECH BOX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART; SM00487; DEXDC;
InterPro; IPR002522; H
InterPro; IPR002521; H
InterPro; IPR002531; H
InterPro; IPR002531; H
InterPro; IPR004109; H
InterPro; IPR004109; H
                                                                                                                                                                  InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3D-structure
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
```

Ø

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license eigenement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELDE FOOTENTIAL).
MAJOR ENVELOE FOOTENTIAL).
MAJOR ENVELOE FOOTENTIAL).
NOINTENDE CICLORC.
N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genotypes.",
Virology 188:331-341(1992)

'Virology 188:331-341(1992)

'I PUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

'I SUBBUNIT: THE VIRICH OF THIS VIRUS IS A NICLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ВУ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01549; HCV_env; 1.
Prodm; PF01560; HCV_NS1; 1.
Prodm; PF01560; HCV_NS1; 1.
Prodm; PF01560; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein; Transmembrane; Nonstructural protein.
INIT_MET : REMOVED FROM CAFSID PROTEIN C I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D10074; BAA00968.1; -. InterPro; IPR002522; HCV_caps1d. InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
CHAIN
CHAIN
CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
HIDDAY SARARAN SARARAN
```

ö

Gaps

ö

Indels

;

0; Mismatches

1 Similarity 100. 28; Conservative

Best Local Matches

```
NCBI_TaxID=31644;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               192
384
347
196
209
233
418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ransmembrane;
                                                                                                                                                          POLG_HCVHK
Q01403;
                                                                                                                                                                                                                                                                                Hepacivirus
 NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
CHAIN
CHAIN
TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                  POLG_HCVHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                    .
P
                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE.
CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN ( POTEWIAL).
MATRIX PROTEIN ( POTEWIAL).
MAJOR ENVELOPE PROTEIN E (POTEWIAL).
MAJOR ENVELOPE PROTEIN E (POTEWIAL).
                                                                                                                                                                                                                                                                                                           Abe K., Inchauspe G., Fujisawa K.;
Abe K., Inchauspe G., Fujisawa K.;
Formonic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                 01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment).
Hepatitis C virus (isolate HCV-476) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                             ;
                       Length 513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Coat protein; Envelope protein;
                                             Indels
55704 MW; 943F31E3514CDEF3 CRC64;
                     Watch 94.2%; Score 145; DB 1; LA Local Similarity 96.4%; Pred. No. 4.1e-14; es 27; Conservative 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BY SIMILARITY
N-LINKED (GLC
                                                                   1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                             PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Transmembrane; Nonstructural protein,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D10688; BAA01530.1; -.
InterPro; IPR002522; HCV_capeld.
InterPro; IPR002513; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
IPfam; PF01543; HCV_capes.d.; I.
Pfam; PF01559; HCV_core; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Co
                                                                                                                                               STANDARD;
513 AA;
                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                            NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF01560;
                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                               POLG_HCVH4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                       Ouery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INIT_MET
                                                                                                                                                           001404;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                          RESULT 3
POLG_HCVH4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                            Matches
 S
                                                                                       g
                                                                                                                                                 ð,
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED. Outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-9310930; Pubmed-1383400;
Abe K., Inchauspe G., Fujisawa K.;
Abe K., Inchauspe G., Fujisawa K.;
Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2739(1992).
I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIDOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genome polyprotein [Contains: Capsid protein C (Core protein) (P22); Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1)! (Fragment).

Hepatitis C virus (Isolate HCV-KF) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nonstructural protein.
                                                                                     Score 145; DB 1; Length 520; Pred. No. 4.1e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CAPSID PROTEIN C (POTENTIAL). MATRIX PROTEIN (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                                                                                                     1; Indels
520 520
520 AA; 56499 MW; AA135246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BY SIMILARITY.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                      520 AA
                                                                                                                                                  0; Mismatches
                                                                                                                                                                                                              1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                           InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR00251; HCV_core.
InterPro; IPR00251; HCV_env.
Pfam; PF01543; HCV_capsid; I.
Pfam; PF01542; HCV_core; I.
Pfam; PF01549; HCV_core; I.
Pfam; PF01560; HCV_env; I.
ProDom; PD186062; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D10687; BAA01529.1; -. PIR; JQ1925; JQ1925.
                                                                                  Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
```

us-09-491-146a-25.rsp

```
ò
                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CELLULAR AMINOPEPTIDAGE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/F2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-9233022; Pubmed-1314459;
A Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
Okamoto H., Fukuda S., Tsuda F., Mishiro S.;
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
Telli-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes.;
Telli-length Sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct c genotypes.;
Telli-length Replace is comparative study of four distinct The Virology 188 331-341(1992).

C --- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN MAND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                    01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation protein C (Core protein); Matrix
17-OCT-2001 (Rel. 40, Matrix
18-OCT-2001 (Rel. 40, Matrix
18-OC
                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     æ
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REMOVED FROM CAPSID PROTEIN C BY
                                                                                                                      Score 145; DB 1; Length 520;
Pred. No. 4.1e-14;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Coat protein; Envelope protein;
                                                                                   1D2BD0A6FF27349B
                                                                                                                                                                                                                                                                                                                                                     Ž
                                                                                                                                                                                                                                                                                                                                                     737
                                                                                                                                                                                                        Pfam: PF01560; HCV_NS1; 1.
Probon: PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Cost protein Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Interpro; IPR00252; HCV_capsid.
Interpro; IPR00252; HCV_core.
Interpro; IPR002519; HCV_env.
Interpro; IPR002531; HCV_env.
Pfam; PP01543; HCV_capsid; 1.
Pfam; PP01543; HCV_capsid; 1.
Pfam; PP01543; HCV_core; 1.
424 N-431 N-449 N-520 56476 MW;
                                                                                                                      ch 94.2%;
l Similarity 96.4%;
27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; D10075; BAA00969.1; -.
424
431
449
520
520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11112;
                                                                                                                                                                                                                                                                                                                            POLG_HCVJ5
ID POLG_HCVJ5
AC P27960;
                                        CARBOHYD
NON_TER
SEQUENCE
CARBOHYD
CARBOHYD
                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INIT_MET
                                                                                                                                             Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                               Matches
    FF FF S
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                     E SE
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 24). Last sequence update)
Genome polyprotein [Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
Hepatitis C virus (isolate HC-J7) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
NCBI_TAXID=11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4
                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                              (POTENTIAL).
                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 737;
369 POTENTIAL.

196 N-LINKED (GLCNAC. . . ) (POTEN CONTROL CON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 145; DB 1;
Pred. No. 6.1e-14;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      737 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_sol.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94.28;
ilarity 96.48;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
         Best_Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVJ7
P27961;
                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
SEQUENCE
                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 6
POLG_HCVJ7
```

S

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-By Step-2003 (Rel. 42, Last annoctation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (Lsolate BR) (HCV)
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                         CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; **Structure and organization of the hepatitis C virus genome isolated
pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01550; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
ProDom; PO186062; HCV_NS1; 1.
ProDom; PO186062; HCV_NS1; 1.
ProDom; PO186062; HCV_NS1; 1.
ProDyprotein; Charles protein; Envelope protein;
Pransmembrane; Nonstructural protein.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-96235224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 737;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 145; DB 1; Length 73
Pred. No. 6.1e-14;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                           CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N LLINKED
N LLIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE-91140698; Pubmed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 94.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 Similarity 96.4 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       233
299
305
417
423
430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                737
737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local :
                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 7
POLG_HCVBK
                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
      DR NO DOR NO DOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 5 E
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                    -i- FUNCTION: THE SWALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-i- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyproteain, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
-i- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                  (RNA)(N).
SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                   Ø
                                                                             LOVE R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
The crystal structure of hepatitis C virus NS3 proteinase reveals
trypsin-like fold and a structural zinc binding site.";
cell 87:331-342(1996).
                                                                                                                                                                         X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.
MEDLINE-8927846, PubMed-9568891.
Yan Y. Li Y. Munshi S., Sardana V., Cole J.L., Sardana M.,
Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
Complex of NS3 protease and NS4A peptide of BK strain hepatitis C virus: a 2.2-A resolution structure in a hexagonal crystal form.",
Protein Sci. 7:837-847(1998).
phosphorylation mediated by CAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE=97015088; PubMed-8861916;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR007095; RNA_pol_DS_PS.
IPR007094; RNA_pol_PSvir.
543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_RGRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1A10; 25-MAR-98.
1JXP; 14-JAN-98.
1NS3; 08-APR-98.
1C2P; 15-NOV-00.
1CSJ; 08-NOV-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1GX5; 09-APR-02.
1GX6; 10-APR-02.
1QUV; 26-JUN-00.
8OHM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GNWVTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF01543;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                    -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PDB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PDB
```

HCV\_NS4a;

PF01006;

Pfam;

327189 MW; F8422D5ECCFDFD9C CRC64;

ž

1188 1197 1202 1204 1688

```
PIR; A39253; GNWVCJ
 11187
11189
11198
1203
1680
                                                                                                                                                                                 Hepacivirus
                                                                                                    POLG_HCVJA
            HELIX
TURN
STRAND
SEQUENCE
 TURN
                                                                                                         P26662;
                                                                                         RESULT 8
POLG_HCVJA
                                                                                                     FT
FT
FT
SO
SO
                                                              à
                                                      CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1, POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF01506; HCV_NS5a; 1.
SMART; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Cat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                   REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                               (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                 (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                             POTENTIAL.
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM
ATP (POTENTIAL).
                                                                                                                                              N-LINKED (GLCNAC.
                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                     N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
                                                                                                                                         DECH BOX.
                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                        LINKED
                                                                                                              081
085
087
097
103
                                                                                                                                                                                                                                                                                                   10082
10086
10090
10090
111009
111120
11120
11130
11130
11168
11168
11168
                                                                                                                                                                                                                                            25240
2529
2788
2788
1031
1039
1050
1059
1068
                                                                                                                                              196
209
234
250
305
417
                                            3D-structure,
INIT_MET
                                                                                                                                                                                CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                              TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                              ACT_SITE
NP_BIND
                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                          CARBOHYD
                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                        STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                         STRAND
STRAND
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                     STRAND
STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                       STRAND
                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                             STRAND
                                                                                                                                                                                                                                                                                   STRAND
                                                                                                                                                                                                                                                                                             STRAND
                                                                                                                                                                                                                                                                                                              STRAND
                                                                       CHAIN
                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                   HELIX
                                                             CHAIN
                                                                   CHAIN
                                                                                              CHAIN
                                                                                                    CHAIN
                                                                                                         CHAIN
                                                                                                                                                                                                                                                                   HELIX
                                                                                                                                                                                                                                                                                        TURN
                                                                                                                                                                                                                                                                                                        TURN
                                                                                                                                                                                                                                                                                                                    TURN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                              ö
                                                                                                                                                                                                                                                                                                                                                01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein [Contains capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NAI); Protein P7; Nonstructural protein NS2 (GP3) (GP34) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NSAB (P65); Nonstructural protein NSAB (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (Isolate Japanese) (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           {RNA}(N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDILINE-91088550; PubMed-2175903;
Kato N., Hilkata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Modecular cloning of the human hepatitis C virus genome from
"Japanese patients with non-B hepatitis";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990);
     Length 3010;
                                                           Indels
     Score 145; DB 1;
Pred. No. 2.9e-13;
                                                                                                                                                                                                                                                                                                   PRT; 3010 AA.
                                                        0; Mismatches
                                                                                                                 1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                          EMBL; D90208; BAA14233.1; -.
     94.28;
Query Match 94.2
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISCUSSION OF SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCB1_TaxID=11116;
```

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Transmembrane; Nostructural Fransferase; RNA-directed RNA polymerase;

W Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Helicase; ATP-binding;
TRMAVED FROM CAPSID PROTEIN C BY THE
CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).

MATRIX PROTEIN (POTENTIAL).

MATRIX PROTEIN NSI (POTENTIAL).

MATRIX PROTEIN NSI (POTENTIAL).

MATRIX PROTEIN NSI (POTENTIAL).

TO CHAIN 1007 1615 PROTEINCTURAL PROTEIN NSI (POTENTIAL).

TO CHAIN 1616 1863 2013 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).

TO CHAIN 1863 2013 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).

TO CHAIN 1863 2013 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).

TO CHAIN 2014 3010 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RELAY SYSTEM (BY
RELAY SYSTEM (BY
RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AA993794F46DB185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -LINKED (
                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001650; Helicase.
InterPro; IPR001650; RNA_pol_DS_PS.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR001095; RNA_pol_DS_PI.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01550; HCV_Core; 1.
Pfam; PF01550; HCV_NS3; 1.
Pfam; PF01005; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01005; HCV_NS4; 1.
Pfam; PF01066; HCV_NS4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHARGE
                                                                                                                                                                                                                                                                                                       HCV_NS5a.
HCV_RdRP.
Helicase_C
                                                                                                                                                                                                    HCV_NS3.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMART; SM00487; DEXDC;
                                                                    InterPro; IPR001410; D
InterPro; IPR002522; H
InterPro; IPR002521; H
InterPro; IPR002519; H
InterPro; IPR002519; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1165
                                                                                                                                                                                                                                                                                                  InterPro; IPR002868;
InterPro; IPR002166;
InterPro; IPR001650;
                                                                                                                                                                                                                            InterPro; IPR004109;
InterPro; IPR000745;
                                                                                                                                                                                                                                                                                InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ¥.
                       MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SITE
DR 20 DR 20
```

Gaps

ö

Length 3010;

Score 145; DB 1; Length 30 Pred. No. 2.9e-13; 0; Mismatches 1; Indels

94.2%; 96.4%;

Conservative

27;

Similarity

Query Match Best Local S Matches 27

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation between the Swinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                            01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
66-00me polyprotein (Contains: Capsid protein C (Core protein) (P22);
67-00me polyprotein El (GP32) (GP35); Envelope glycoprotein El (GP68) (GP76) (MS1); Protein P7: Nonstructural protein NS2 (P21)
6C 3.4.99-); Protease/helicase NS3 (P70) (Hepacivirin)
6C 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A., Murcko, M.A., Lin C., Caron P.R.; "Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       H.V., Weber P.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         [RNA](N).

-1 SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.

-1 PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.

-1 SIMILARITY: THE NS? PROTEASE BELONGS TO PEPTIDASE FAMILY 199.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE-92052256; Pubmed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V.
"Structure of the hepatitis C virus RNA helicase domain."
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
                                                                                                                                                                                               PRT; 3011 AA
                             PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-98154321; PubMed-9493270;
                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       of unwinding.";
Structure 6:89-100(1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                            POLG_HCVH
P27958;
```

```
Gaps
 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1627
1636
1652
A: 327142 MW; 772CBB29CCD94753 CRC64;
     ??
. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 145; DB 1; I
Pred. No. 2.9e-13;
0; Mismatches 1;
 N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
 1317
1323
1336
1343
1352
1360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
POLG_HCVJ6
ID POLG_HCVJ6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HELIX
TURND
STRAND
STRAND
STRAND
HELIX
HELIX
HELIX
TURN
HELIX
TURN
HELIX
TURN
HELIX
STRAND
 CARBOHYD
CARBOHYD
                                 STRAND
TURN
HELIX
TURN
HELIX
TURN
HELIX
TURN
STRAND
STRAND
STRAND
STRAND
TURN
HELIX
TURN
HELIX
TURN
TURN
STRAND
TURN
HELIX
TURN
HELIX
TURN
TURN
TURN
TURN
TURN
TURN
TURN
STRAND
STRAND
STRAND
TURN
STRAND
TURN
STRAND
STRAND
TURN
STRAND
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
 (See http://www.isb-sib.ch/announce/
     DR PREL, M67463, AAA45341; -
DR PRE, A36814 GWWVCH.
DR PDB; IAIN, 10-PEB-99.
DR MERCPS; U39-001; -
DR TANNERAC; TQ4155; -
DR TANNERAC; TQ4155; -
DR INCEPTO; IPRO02512; HCV_core.
DR INCEPTO; IPRO02512; HCV_core.
DR INCEPTO; IPRO02513; HCV_MS2.
DR INCEPTO; IPRO02519; HCV_MS3.
DR INCEPTO; IPRO02519; HCV_MS3.
DR INCEPTO; IPRO02169; HCV_MS3.
DR INCEPTO; IPRO01409; HCV_MS3.
DR INCEPTO; IPRO01609; RNA_DOL_DS_PS.
DR INCEPTO; IPRO01609; RNA_DOL_DS_PS.
DR INCEPTO; IPRO01609; HCV_MS3.
DR PEAM; PFO159; HCV_CORF; 1.
DR PEAM; PFO159; HCV_CORF; 1.
DR PEAM; PFO159; HCV_MS3; 1.
DR PEAM; PFO159; HCV_MS3; 1.
DR PEAM; PFO150; HCV_MS3; 1.
DR PEAM; PFO1001; HCV_MS4; 1.
DR PEAM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REMOVED FROM CAPSID PROTEIN C BY THE CAPSID PROTEIN C.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHARGE RELAY SYSTEM (BY SI ATP (POTENTIAL).

DECH BOX.

N-LINKED (GLCNAC...) (PO N-LINKED (GLCNAC...)) (PO N-LINKED (GLCNAC...) (PO
:itles requires a license agreement (
send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CCHAIN
CCHAIN
CCHAIN
CCHAIN
CCHAIN
CCHAIN
CCHAIN
CCHAIN
CCHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
CARBOHYD
CAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
```

```
POLYPOTOTELIN, Glycoprotein; Transferase; RNA-directed RNA polymerase; W Polyprotein; Glycoprotein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolases; Serine protease.
Transmembrane; Nonstructural protein; Hydrolases; Serine protease.
Transmembrane; Nonstructural protein; Grysling; PRWOVED FROM CAPSID PROTEIN C FOTEWITAL).
TCHAIN 116 191 MAJOR ENVELOPE PROTEIN E FOOTEWITAL).
TCHAIN 192 383 MAJOR ENVELOPE PROTEIN NS (POTEWITAL).
TCHAIN 184 733 NONSTRUCTURAL PROTEIN NS (POTEWITAL).
TCHAIN 1670 1010 PROTEASE/HELICASE NS3 (POTEWITAL).
TCHAIN 1670 1010 PROTEASE/HELICASE NS3 (POTEWITAL).
TCHAIN 1670 1010 NONSTRUCTURAL PROTEIN NS48 (POTEWITAL).
TCHAIN 1670 1010 NONSTRUCTURAL PROTEIN NS48 (POTEWITAL).
TRANSMEM 347 369 POTEWITAL.
TCHAIN 1011 1111 CHARGE RELAX SYSTEM (BY SIMILARITY).
TCT_SITE 1016 1111 CHARGE RELAX SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-406-1992 (Rel. 23, Created)
01-406-1992 (Rel. 23, Last sequence update)
01-406-1992 (Rel. 23, Last sequence update)
01-406-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains Capsid protein C (Core protein) (P22);
Envelope 91ycoprotein E1 (GP32) (GP35); Envelope 91ycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/hellcase NS3 (P70) (Repartin)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS58 (P65) (P70) (RNA-directed RNA pollymerase) (EC 2.7.7.48)].
NS58 (P65) (P70) (RNA-directed RNA pollymerase) (EC 2.7.7.48)].
Hepatitis C virus (Laslate HC-J8) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 145; DB 1;
Pred. No. 2.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL).
DBCH BOX.
DBCH BOX.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PKPQRKTKRNTYRRPQDVKFPGGGQIVG
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SWART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               94.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus.
NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2038
2811
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Sim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLG_HCVJ8
P26661;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVJ8
1D POLG_H
AC POLG_H
DT 01-AUG
DT 01-AUG
DT 02-PEB
DE GENOME
DE GF089
DE (GF08)
DE (GF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loc
Matches
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=92044440; PubMed=1658196;
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
Adhato A., Miyakawa Y., Mayumi M.;
Machida A., Miyakawa Y., Mayumi M.;
Mucleotide sequence of the genomic RNA of hepatitis C virus isolated
from a human carrier: comparison with reported isolates for conserved
and divergent regions. "10. Gen. Virol. 72:8697-2704(1991).
J. Gen. Virol. 72:8697-2704(1991).
J. Gen. Virol. 72:8697-2704(1991).
C -!- FUNCTION: THE SMALL PROTEINS NSZA, NSZB, NSZA AND NSZB ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NSJ AND NSS MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NSJ AND NSS MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C -!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral presented commonly with Asp or Glu in the P6
precursor polyprotein, commonly with all np 1.
C -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (N81); Protean P7: Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NSA (P45); Nonstructural protein NSA (P65); NONSTRUCTURAL NSA (P65); NONSTRU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (RNA)(N).
SUBDUIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Interpro; IPR002166; HCV_RGRP.
Interpro; IPR001650; Helicase_C.
Interpro; IPR007095; RNA_pol_bS_PS.
Interpro; IPR007094; RNA_pol_bSvir.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D00944; BAA00792.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PF01542; HCV_core;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR, JQ1303, JQ1303.
. HSSP, P27958, 1HEI.
MEROPS, S29.001; -.
MEROPS, U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002518;
InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
```

ö

Gaps

ö

Pfam; Pfam;

Pfam; Pfam; Pfam;

Pfam; Pfam;

SIMILARITY). SIMILARITY).

(BY (BY (BY

(POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL)

(GLCNAC GLCNAC (GLCNAC (GLCNAC (GLCNAC (GLCNAC (GLCNAC

(POTENTIAL)

GLCNAC GLCNAC (GLCNAC (POTENTIAL

(POTENTIAL) (POTENTIAL) (POTENTIAL)

GLCNAC

(POTENTIAL) (POTENTIAL)

> (GLCNAC (GLCNAC

```
RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                  1A173E7E3381FD1A CRC64;
                          CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                         (GLCNAC.
                                                                  ATP (POTENTIAL)
                                                                                                                             N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                       N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                             DECH BOX
                                                                                                                                                                                                                                                                                                                                                                         94.28;
                                                                                                                                                                                                                                                                                                                                                 330177
                                                                                                                                                                                                                                                                                                                                                                                      Best_Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                     2811
                                                                                                                                                                                                                                                                                               1091 100
2038 200
2359 230
2811 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCVTW
P29846;
                           ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
   CHAIN
TRANSMEM
                                                                                         CARBOHYD
                                                                                                       CARBOHYD
                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
POLG_HCVTW
   This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                           between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CAPSID PROTEIN C (POTENTIAL).
MACHRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                       PROJECTION OF THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HUNCHION: THE SMALL PROTEINS NS2A, NS2B, NS4B AND NS4B ARE HYDROPHORIC, SUGGESTING A POSSTBLE MEMBRANE-RELATED FUNCTION.

1. CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

1. CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
SEQUENCE FROM N.A.
MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
Full-length sequence of a hepatitis C virus genome having poor
homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Hallcase; APP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                               SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                       SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410; DEAD.
InterPro; IPR001410; DEAD.
InterPro; IPR002521; HCV_core.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR00219; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR001868; HCV_NS5a.
InterPro; IPR002166; HCV_RSAP.
InterPro; IPR002166; HCV_RSAP.
InterPro; IPR007095; RNA_POI_DS_PS.
InterPro; IPR007094; RNA_POI_DS_PS.
InterPro; IPR007094; RNA_POI_DS_PS.
InterPro; IPR01643; HCV_CORE!
                                                                                                                                                                                                                                                                                                                                                                                 EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF01506; HCV_NS5a; 1
Pfam; PF00998; Viral RdRP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                          HSSP, P27958; 1HEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   116
192
384
384
734
1011
1620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
```

```
01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 42, Last sequence update)
16-GP69 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22..); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-92230206; PubMed=1314449;
Chen P.J. Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
The Taiwanese hepatitis C virus genome: sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA.";
Virology 188:102-113(1992).
-!-FUNCTION: THE SMALL PROYEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE WEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-!-CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the vira precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.
-!-CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                           ő
    Length 3033;
                                                                                       1; Indels
Score 145; DB 1;
Pred. No. 2.9e-13;
                                                                                       0; Mismatches
                                                                                                                                                                     1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
```

us-09-491-146a-25.rsp

```
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                entities requires a license agreement (See http://www.lsb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-AUG-1992 (Rel. 23, Lest sequence update)
15-SEP-2003 (Rel. 42, Lest annotation update)
16-SEP-2003 (Rel. 42, Lest annotation Elevation Elevatory (GP69 (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
16-SEP-2003 (Rel. 42, Lest annotation Elevatory (Rel. 42, Lest annotation)
16-SEP-2003 (Rel. 42, Lest annotation Elevatory (Rel. 42, Lest annotation)
16-SEP-2003 (Rel. 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-91172826; PubMed-1848704;
Choo Q.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,
Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,
Bradley D.W., Kuo G., Houghton M.;
Genetic organization and diversity of the hepatitis C virus.";
Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1991).
-!-FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-!-CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
-!-CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                           (POTENTIAL)
                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC. . .) (POTENTIAL)
WW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                          POTENTIAL)
                                                                         POTENTIAL)
                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 1; Length su. Pred. No. 3.3e-12;
                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                           (GLCNAC
                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6 KPQRKTKRNTNRRPQDVKFPGGGQIVG 32
         N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 KPORKTKRNTYRRPODVKFPGGGGIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-AUG-1992 (Rel. 23, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                   89.68;
96.38;
                                                                                                                                                                                                                                                                                                                                                                                            327047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11104;
                                                                                                                                                                                   623
645
2041
2077
2254
2529
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ma.
Local Sin.
26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCV1
                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                     CARBOHYD
                                                                                                                               CARBOHYD
                                                                                                                                                              CARBOHYD
                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCV1
            STATESTATES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DAH PUDU BERMANANAN PUDU BERMAN PUDU BERMAN PUDU BERMANAN 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBC outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CELLULAR AMINOPEPTIDASE.

ONE PROTEIN (POTENTIAL).

MAJOR ENVELOPE PROTEIN E (POTENTIAL).

MAJOR ENVELOPE PROTEIN E (POTENTIAL).

NONSTRUCTURAL PROTEIN NSJ (EO POTENTIAL).

PROTEASE/HELICASE NSJ (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NONSTRUCTURAL PROTEIN NGAA (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                               SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_MS1.
InterPro; IPR002519; HCV_MS2.
InterPro; IPR004109; HCV_MS3.
InterPro; IPR001490; HCV_MS4a.
InterPro; IPR001490; HCV_MS4a.
InterPro; IPR001490; HCV_MS4b.
InterPro; IPR001699; HCV_MS5a.
InterPro; IPR001699; RCV_MS5a.
InterPro; IPR001995; RNA_PO1_DS_FS.
InterPro; IPR007094; RNA_PO1_DS_FS.
InterPro; IPR007094; RNA_PO1_PS_VIT.
Pfam; PP01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                             EMBL; M84754; -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF01538: HCV_NS2; 1.
PF02907: HCV_NS3; 1.
PF01006: HCV_NS4s; 1.
PF01001: HCV_NS4s; 1.
PF01001: HCV_NS4s; 1.
PF01506: HCV_NS4s; 1.
PF00291: Helicase_C; 1.
PF00998: Viral_RdRP; 1.
Om: PD166062: HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF01539; HCV_env;
PF01560; HCV_NS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     369
11083
11165
11165
1234
1234
233
233
4417
423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1006
1615
1862
2013
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                       PDB; 1N64; 25-FEB-03
PDB; 1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                            PIR; A40244; GNWVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01542;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3D-structure,
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
```

EMBL; M62321; AAA45676.1; -.

```
Leduc R., Molloy S.S., Thoras B.A., Thomas G.;
Leduc R., Molloy S.S., Thoras B.A., Thomas G.;
"Activation of human furin precursor processing endoprotease occurs by an intramolecular autoproteolytic cleavage.";
J. Blol. Chem. 267:14304-14308 (1992).
-1- FUNCTION: FURIN IS LIKELY TO REPRESENT THE UBIQUITOUS ENDOPROTEASE ACTIVITY WITHIN CONSTITUTIVE SECRETORY PATHWAYS AND CAPABLE OF CLEAVAGE AT THE RK(K/R)R CONSENSUS MOTIF.
-1- CATALYTIC ACTIVITY: Release of mature proteins from their proproteins by cleavage of Agg-Xaa-Yaa-Agg-1-Zaa bonds, where Xaa can be any amino acid and Yaa is Arg or Lys. Releases albumin, complement component C3 and von Willebrand factor from their
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 296-794 FROM N.A. MEDIAGE SEQUENCE OF 296-794 FROM N.A. MEDIATE-87053585; PubMed-3023061; Reobrock A.J.M., Schalken J.A., Leunissen J.A.M., Onnekink C., Bloemers H.P.J., van de Ven W.J.M.; "Evolutionary conserved close linkage of the c-fes/fps proto-oncogene and genetic sequences encoding a receptor-like protein."; EMBO J. 5:2197-2202(1986).
                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3D-STRUCTURE MODELING OF CATALYTIC DOMAIN.
MEDLINE-94291619; PubMed-8020465;
Siezen R.J., Creemers J.W.M., van de Ven W.J.M.;
"Homology modelling of the catalytic domain of human furin. A model for the eukaryotic subtilisin-like proprotein convertases.";
Eur. J. Blochem. 222:255-266(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "CDNA and gene structure for a human subtilisin-like protease with cleavage specificity for paired basic amino acid residues."; DNA Cell Biol. 10:319-328(1991).
                                                                                                                                                                                                                                                                                                                                           Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                           FURI_HUMAN STANDARD; PRT; 794 AA.
P09958: 014336;
01-MAR-1989 (Rel. 14, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
15-SEP-2003 (Rel. 42, Last annotation update)
15-SEP-2009 (Rel. 42, Last annotation update)
16-SEP-2009 (Rel. 42, Last annotation update)
17-SEP-2009 (Rel. 42, Last annotation update)
18-SEP-2009 (Rel. 42, Last annotation update)
18-SEP-2009 (Rel. 43, Last annotation update)
19-SEP-2009 (Rel. 44, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                            TISSUE-Blood;
MEDLINE-90175002; PubMed-2408021;
wan den Ouweland A.M.W., van Duljnhoven H.L.P., Keizer G.D.,
Dorssers L.C.J., van de Ven W.J.M.;
Structural homology between the human fur gene product and the
subtilisin-like protease encoded by yeast KEX2.";
Nucleic Acids Res. 18:664-664(1990).
                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
MEDLINE-91321735; Pubmed-1713771;
Barr P.J., Mason O.B., Landsberg K.E., Wong P.A., Kiefer M.C.,
                                     Length 3011;
                       ; DB 1; Lens. 4 6e-12; Indels
   327197 MW; 65F8C9447FCE5AF9 CRC64;
                                 Score 137; DB Pred. No. 4.6e-1; Mismatches
                                                                                                  1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                 MEDLINE-92332543; PubMed-1629222;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COFACTOR: CALCIUM-DEPENDENT
                                   89.0%;
                                                 Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       respective precursors
     3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Brake A.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROCESSING
     SEQUENCE
                                     Query Match
                                                                                                                                                                                                 FURI_HUMAN
                                                                                                                                                                                RESULT 14
                                                                                                                            ద
                                                                                                                                                                                                                  å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN NSI,PEZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI,PEZ (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS24 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
(GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                            InterPro; IPR00156; HCV_RGRP.
InterPro; IPR00150; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
Pfam; PF01543; HCV_capsid; I.
Pfam; PF01542; HCV_core.
                                                                                           _capsid.
                                                                                                                             HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                           /_core.
                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01056; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1006
1615
1862
2013
3011
369
1107
11165
1237
1319
196
209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SMART; SM00487; DEXDC;
             PDB; 1A1V; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                           IPR002868; H
IPR002166; H
                                                                              InterPro; IPR001410;
                                                                                               InterPro; IPR002522;
                                                                                                           InterPro; IPR002521;
A39166; GNWVC3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   347
1083
11107
11165
11165
1316
1316
196
209
209
234
305
                                                                                                                                                                                                                                                                                                                                                         PF01560;
                                                                                                                                                                                                                                                                                                                                                                           PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3D-structure.
INIT_MET
                                                                                                                                                          InterPro;
InterPro;
                                                                                                                             InterPro;
                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
```

```
-1- SUBCELLULAR LOCATION: SEEMS TO BE LOCALIZED INTRACELLULARLY TO THE TRANS GOLGI NETWORK. PROPEPTIDE CLEAVAGE IS A PREREQUISITE FOR EXIT OF FURIN MOLECULES OUT OF THE ENDOPLASMIC RETICULUM (ER).

SECOND CLEAVAGE IN THE PROPEPTIDE OCCUR IN THE TRANS GOLGI NETWORK (TGN), WHICH IS FOLLOWED BY THE RELEASE OF THE PROPEPTIDE BOUND TO FURIN AND THE ACTIVATION OF FURIN.

-1- TISSUE SPECIFICITY: SEEMS TO BE EXPRESSED UBIQUITOUSLY.

LOCALIZATION AND RECYCLING FROM THE CELL SURFACE.

-1- DOMAIN: CONTAINS A CYTOPLASMIC DOMAIN RESPONSIBLE FOR ITS TGN LOCALIZATION AND RECYCLING FROM THE CELL SURFACE.

-1- PTW: THE PROPEPTIDE IS AUTOCATALYTICALLY REMOVED THROUGH AN INTRAMOLECULAR CLEAVAGE PROBABLY IN THE BNOPLASMIC RETICULUM (ER). IN THE TGN THE SECOND CLEAVAGE IN THE PROPEPTIDE COULD LEAD
                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics. The Flutte. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
ENZYME REGULATION: COULD BE INHIBITED BY THE NOT SECONDLY CLEAVED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS00136; SUBTILASE_ASP; 1.
PROSITE; PS00137; SUBTILASE_HIS; 1.
PROSITE; PS00138; SUBTILASE_SER; 1.
Hydrolase; Serine protease; Transmembrane; Glycoprotein; Signal; Zymogen; Calcium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY)
SIMILARITY)
                                                                                                                                                                                                                                                               SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY. SIMILARITY: Contains 1 homo B/P domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC. . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CLEAVAGE (SECOND AUTO').
CLEAVAGE (FIRST AUTO').
CLEL ATTACHMENT SITE (POTENTIAL).
CELL SURFACE SIGNAL.
TRANS GOLGI NETWORK SIGNAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC. . .) (POTENTIAL) (GLCNAC. . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10C44DD5892EF85D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       136950, ... Golgi apparatus, TAS. GO:0005794; C:Golgi apparatus; TAS. GO:0004276; F:furin activity; TAS. GO:0007267; P:cell.cell signaling; TAS. GO:0006508; P:proteolysis and peptidolysis; TAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro: IPR006212; Furin_repeat.
InterPro: IPR002884; P. domain.
InterPro: IPR00209; Peptidase_S8.
Pfam; PF01483; P. Proprotein; PARTIAL.
Pfam; PR001483; P. Proprotein; PARTIAL.
ProDom: PR00082; Peptidase_S8; 1.
ProDom: PD000717; P. domain; 1.
SNART; SM00261; FU; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CYS-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FURIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          86678 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; X17094; CAA34948.1; -. EMBL; X04329; CAA27860.1; -. EMBL; A06939; CAA00605.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS; S08.071; -. Genew; HGNC:8568; FURIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                107
500
762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1MPT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; A39552; KXHUF.
HSSP; Q99405; 1MPT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          794 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROPEP
 <del>.</del>
                                   -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8888888888888888888888
```

DB 1; Length 794;

Score 60; DB 1, Pred. No. 0.41;

39.0%; 55.6%;

Best Local Similarity

Query Match

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       - I SUBCELLUIAR LOCATION: SEEMS TO BE LOCALIZED INTRACELLULARLY TO THE TRANS GOLGI NETWORK. PROPEPTIDE CLEAVAGE IS A PREREQUISITE FOR EXIT OF FURIN MOLECULES OUT OF THE ENDOPLASMIC RETICULUM (ER).

SECOND CLEAVAGE IN THE PROPEPTIDE OCCUR IN THE TRANS GOLGI NETWORK (TGN), WHICH IS FOLLOWED BY THE RELEASE OF THE PROPEPTIDE BOUND TO FURIN.

- I TISSUE SPECIFICITY: SEEMS TO BE EXPRESSED UBIQUITOUSLY.

- I DOMAIN: CONTAINS A CYTOPLASMIC DOMAIN RESPONSIBLE FOR ITS TGN LOCALIZATION AND RECYCLING FROM THE CELL SURFACE.

- PTM: THE PROPEPTIDE IS AUTOCATALIYTICALLY REMOVED THROUGH AN INTRAMOLECULAR CLEAVAGE PROBABLY IN THE ENDOPLASMIC RETICULUM (ER). IN THE TGN THE SECOND CLEAVAGE IN THE PROPEPTIDE COULD LEAD TO THE ALTHURAMOLECULAR CLEAVAGE PROBABLY IN THE PROPEPTIDE COULD LEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Creemers J.W.M., Roebroek A.J.M., van den Ouweland A.M.W.,
Creemers J.W.M., Roebroek A.J.M., van de Ven W.J.M.;
van Duijnhoven H.L.P., van de Ven W.J.M.;
Cloning and functional expression of a 4.3 kbp mouse fur cDNA:
evidence for differential expression.";
Life Sci. Adv. (Mol. Biol.) 11:127-138(1992).
I. FUNCTION: FURIN IS LIKELY TO REPRESENT THE UBIQUITOUS ENDOPROTEASE
ACTIVITY WITHIN CONSTITUTIVE SCRETCRY PATHWAYS AND CAPABLE OF
CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF.
I. CATALYTIC ACTIVITY: Release of mature proteins from their
Proproteins by cleavage of Arg-Xaa-Yaa-Arg--Zaa bonds, where Xaa
can be any amino acid and vaa is Arg or Lys. Releases albumin,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BY SIMILARITY).
INHIBITED BY THE NOT SECONDLY CLEAVED
    Gaps
                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    complement component C3 and von Willebrand factor from their respective precursors.
COFACTOR CALCIUP-BEPRINGER (BY SIMILARITY).
ENZYME REGULATION: COULD BE INHIBITED BY THE NOT SECONDLY CLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY. SIMILARITY: Contains 1 homo B/P domain.
                                                                                                                                                                                                                                                        01-NOV-1991 (Rel. 20, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Furin precursor (EC 34.21.75) (Paired basic amino acid residue cleaving enzyme) (PACE) (Dibasic processing enzyme).
  ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Structure and expression of mouse furin, a yeast Kex2-related protease. Lack of processing of coexpressed prorenin in GH4Cl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hatsuzawa K., Hosaka M., Nakagawa T., Nagase M., Shoda A.,
Indels
                                                                                                                                                                                                793 AA
    Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                J. Biol. Chem. 265:22075-22078(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                        [1]
SEQUENCE FROM N.A.
MEDLINE-91093035; Pubmed-2266110;
                                                                :|:|||: |: | | |||
102 KRRTKRDVYQEPTDPKFP 119
                                          21
                                                                                                                                                                                                                                      (Rel. 20, Created)
                                          4 ORKTKRNTYRRPODVKFP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Murakami K., Nakayama K.;
  10; Conservative
                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                           cleaving enzyme) (PAC
FURIN OR FUR OR PCSK3
                                                                                                                                                                                                                                                                                                                                                                     musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROPEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE-Liver;
                                                                                                                                                                                                                                      01-NOV-1991
                                                                                                                                                                                                MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            cells.
  Matches
                                                                                                                                                     RESULT 15
                                                                                                                                                                                                                                                                                                                                                                     Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ++
                                                                                  g
                                                                                                                                                                                                à
```

```
| PIR; A23679; XMSEP. | PIR; A23679; XMSEP. | PIR; A23679; XMSEP. | PIR; A23679; XMSEP. | PIRP. | PIRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CYS-RICH.

POTENTIAL.

CHARGE RELAY SYSTEM (BY SIMILARITY).

CHARGE RELAY SYSTEM (BY SIMILARITY).

CHARGE RELAY SYSTEM (BY SIMILARITY).

POTENTIAL.

POTENTIAL.

N-LINKED (GLCNAC...) (POTENTIAL).

N-LINKED (GLCNAC...) (POTENTIAL).

N-LINKED (GLCNAC...) (POTENTIAL).

CLEAVAGE (SECOND AUTO-).

CELAVAGE (FIRST AUTO-).

CELL SURFACE SIGNAL.

TRANS GOLGI NETWORK SIGNAL.

CELL SURFACE SIGNAL.

CELL STAGRMENT SITE (POTENTIAL).

M -> V (IN REF. 2).

M -> V (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 35.7%; Score 55; DB 1; Length 793; Best Local Similarity 50.0%; Pred. No. 2.3; Matches 9; Conservative 4; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BY SIMILARITY, FURIN,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 86804 MW;
EMBL; X54056; CAA37988.1; -. EMBL; L26489; AAA37643.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             793 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SITE
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
DISULFID
DISULFID
CARBOHYD
CARBOHYD
CARBOHYD
SITE
      DR WEEK WEEK WAREN
```

.rch completed: August 7, 2003, 11:19:58 .o time : 5.90909 secs

4 ORKTKRNTYRRPODVKFP 21

q

ò

; 0

us-09-491-146a-25.rspt

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

US-09-491-146A-25 154 1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

830525 seqs, 258052604 residues Searched:

830525 tal number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

sp\_vertebrate:\*
sp\_unclassified:\* sp\_fungi:\*
sp\_human:\*
sp\_invertebrate:\* sp\_rvirus:\* sp\_bacteriap:\* sp\_archeap:\* sp\_organelle:\* sp\_phage:\* sp\_archea:\* sp\_bacteria:\* sp\_plant:\* sp\_mammal:\* p\_virus:\* SPTREMBL\_23:\* sp\_mhc:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

					٠.			٠.	٠.	٠.						•.	•.	
		5	S	S	S	S	S	S	8	S	S	S	S	S	S	S	S	
		titi	titi	nepatitis	nepatitis	nepatitis	nepatitis	nepatitis	nepatitis	titi	titi	nepatitis	nepatitis	nepatitis	nepatitis	nepatitis	epatitis	
	lon	hepatitis	hepatitis	epa	nepa	hepa	nepa	nepa	eba	hepa	hepatitis	hepa	hepa	hepa	hepa	hepa	hepa	
	escription	1		_	75 ]	00	89	41	07				32	83.	14	82	87 1	
	escr	P90364	281831	Md6C	2685	291a	2819	2815	2818	2683	2683	2683	283y	28.37	Q8 jyr4	28qp	28gp	
	ă	_	_	_	_	Š	Č	_	_	~	_		_	_	_	_	Ŭ	
	1																	
		-	_	~	2	0	6	_	7	7	<b>a</b> o	0	~	<u>س</u>	4	ťΩ	7	
		P90364	081831	PXN.	857	1AU(	198	154	180	068307	8300	831	JYS	JYS	28JYR4	38QP85	80P87	
	a	P9	80	6	90	6	8	80	80	90	90	90	80	8	8	8	80	
	DB	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	
	gth	105	106	125	191	010	010	010	109	45	45	45	9	61	73	100	100	
	Query Aatch Length DB					m	m	m										
	ery	0.0	0.0	100.0	0.0	0.0	0.0	0.0	7.4	4.8	4.8	4.8	4.8	4.8	4.8	4.8	4.8	
æ	o R	10	10	10	10	10	ព	10	φ	6	σ	σ	6	Φ	6	6	σ	
	Score	154	154	154	154	154	154	154	150	146	146	146	146	146	146	146	146	
	Sc																	
	ult No.	-	~	٣	4	2	ဖ	7	œ	σ	10	11	12	13	14	15	16	
	Result No.																	

U	O	O	Ų	υ	O	O	U	O	O	O	ပ	U	O	O	O	υ	Ç	Ç	O	υ	O	O	υ	O	υ	O	υ	ပ
hepatitis	hepatitis	hepatitis	hepatitis	hepatitis					hepatitis	hepatitis	hepatitis	hepatitis				-	hepatitis	hepatitis	hepatitis	hepatitis	hepatitis	hepatitis	hepatitis	-				
Q8qp72	08qp71	08qp74	084580	08qp77	Q8qp84	08qp78	08qp75	08qp79	08qp81	Q8qp83	08qp76	08qp86	081340	Q8qmj4	068892	068893	Q8bcx2	Q8bcx1	08bcx0	Q8bcw9	Q8bcw8	Q8bcw7	QBbcw6	Q8bcw4	08qr11	08qm16	08qm18	Q8qrj4
72	71	74	80	77	84	78	7.5	179	181	83	92	98	40	134	26	93	X2	1X1	0 <b>X</b> :	6M:	.w8	W7	W6	W4	151	II6	118	174
<b>080P</b>	080P	080	080P	080P	080P	080P	080P	Q8QP79	080P	080F	080F	080P	0813	080	0688	0688	<b>08BC</b>	Q8BCX1	Q8BCX0	QBBCW9	Q8BCW8	QBBCW7	QBBCW6	QBBCW4	080F	OBOMI6	080	Q8QRJ
12	12	15	12	12	12	12	12	12	12	12	13	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
100	100	100	100	100	100	100	100	100	100	100	100	100	109	114	114	114	119	119	119	119	119	119	119	119	122	122	122	122
	٠.	٠.	٠,	٠,	94.8	•	•	8.46			•			٠	٠			94.8					٠		٠	94.8		94.8
**	₹*	4	4	4	146	❖	4	•	146	4	4	₹	4	4	4	4	4	146	4	4	4	4	4	4	4	146	4	146
17	18	19	20	21	. 22	23	24	25	56	27	28	58	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

## ALIGNMENTS

RESULT 1 P90364		01-MAY-1997	01-MAY-1997	01-OCT-2002 (	Core protein	OS Hepatitis C virus.		OC Hepacivirus.	OX NCBI_TaxID=11103;	RP SEQUENCE FROM N.A.	RA Ohno T., Mizokami M.;	RT "Determination of nine	RT method.";	RL Thesis (1995), Nagoya City University Medical School.			RA Ohno T., Mizokami M.,				RT patients.";	Arch	<u>:</u>				EMBL			NON_TER	FT NON_TER 105 105 SQ SEQUENCE 105 AA; 11
	ARY;	03,		el. 22, L	polyprot		Lve-stran					o		a City Un	ņ	Med-8712	Saleh M	Nour1-A	ation of	pplicatio		1-1113(19	ION OF TH	COPE. THE	COPROTEI	NA (BY SI	19.1;	HCV_caps	psid; 1.	<del>.</del>	11870 MW;
	PRT; 105 AA.	Created)	Last sequence update)	Last annotation update)	(Genome polyprotein) (Fragment).		d viruses, no					pes of hepati	•	iversity Medi		927;	.G., Orito E.	ria K.T., Tok	phylogenetic	n to isolates		. (96)	IS VIRUS IS A	ENVELOPE CON	N E. THE NUCL	MILARITY).		id.			C7BA40B284025A49
	ъ.		update)	n update)	E).		Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;					genotypes of hepatitis C virus using PCR	•	cal School.			, Ohba KI., Wu RR.,	udome S., Williams R.;	"Usefulness and limitation of phylogenetic analysis for hepatitis C	virus core region: application to isolates from Egiptian and Yemeni			SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A	LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:	EOCAPSID IS A COMPLEX OF						25A49 CRC64;

Gaps

ö

```
Songsivilai(S., Dharakul T., Kunkitti R., Thepthai C.;
"Molecular cloning and expression of hepatitis C virus core protein
and production of monoclonal antibodies to the recombinant protein.";
Asian Pac. J. Allergy Immunol. 14:0-0(1996).

-I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOROPRIEN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Lim S.P., Khu Y.L., Hong W.J., Tay A., Ting A.E., Lim S.G., Tan Y.H.; "Identification and molecular characterization of the complete genome of a Singapore isolate of hepatitis C virus: sequence comparison with other strains and phylogenetic analysis."; Virus Genes 23:89-95(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                        Length 125;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 191;
                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              191
20840 MW; 4AAE63444D8329E2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                        Score 154; DB 12;
Pred. No. 1.2e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 154; DB 12;
Pred. No. 1.9e-15;
; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                          191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 3010 AA
                                                                                                 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100.0%; Preq. ...
                                                                                                                                                                                             1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPORKTKRNTYRRPODVKFPGGGQIVG 32
                                                                                                                                                             1 PKPORKTKRNTYRRPODVKFPGGGOIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequ
01-MAR-2003 (TrEMBLrel. 23, Last anno
                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-HCV-S1;
MEDLINE-21440119; PubMed-11556407;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002522; HCV_caps1d.
InterPro; IPR002521; HCV_core.
                                  Query Match 100.0%;
Best Local Similarity 100.0%;
Matches 28; Conservative 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best_Local Similarity 100.0
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. STRAIN-BB51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q91AU0;
01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                          068575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       091AU0
                                                                                                                                                                                                                                                                                                              RESULT 4
Q68575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 5
Q91AU0
                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                          SY TWEEN THE STATE OF THE STATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCOCCOS SERVING SERVI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ô
                                                                     ô
                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepetitis C virus type 2.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIZCORROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MANA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-94351179; Pubmed-7520922;
Nakazawa T., Kato N., Ohkoshi S., Shibuya A., Shimotohno K.;
Nakazawa T., Kato N., Ohkoshi S., Shibuya A., Shimotohno K.;
-1. Bubuxi: 20:623-629(1994).
-1. SUBUXI: THE VIRION OF THIS VIRIOS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMFLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

InterPro: 1PR002522; HCV_capsid.
Pfam: PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ohno T., Mizokami M.; **Determination of nine genotypes of hepatitis C virus using PCR
                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
   Length 105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 106;
                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
(Isolate USA8) genomic RNA (Genome polyprotein) (Fragment).
                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               106 106
106 AA; 12001 MW; 25D0D5414B3EA9DC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     D945CB60661797E3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0%; Score 154; DB 12;
100.0%; Pred. No. 9.9e-16;
ive 0; Mismatches 0;
100.0%; Score 154; DB 12; 100.0%; Pred. No. 9.8e-16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       E2/NS1 protein (Genome polyprotein) (Fragment)
Hepatitis C virus type 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            125 AA
                                                           0; Mismatches
                                                                                                                                                                                                                                                                                                                                       106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTYRRPQDVKFPGGGGIVG 28
                                                                                                                                                       Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; D49465; BAA08439.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  125 AA; 14220 MW;
                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-OCT-2002 (TrEMBLrel. 22,
                                                           28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity 100 les 28; Conservative
                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-40271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein.
NON_TER 10
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-USA8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
SEQUENCE 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                081831;
                                                                                                                                                                                                                                                                                                                                    081831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9PXN3
                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TSULT 3
                                                                                                                                                                                                                                                                     RESULT 081831 ID 088 AC 088 01 DT 01 DT 01 DT 01 DT 01 DE 01 DE 01 DE 01 DE 01 DE 02 VIEL OC VIEL DE 02 HE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

ô

Gaps

```
STRAIN-JT';

MEDLINE-92295714; PubMed-1318627;

Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J., Nakazawa T.,

Hijikata M., Ishimura Y., Shimotohno K.;

"Moleoular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among infected individuals.";
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE; PSS0507; RDRP_POSITIVE; 1.
PROSITE; PSS0521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3010 AA; 327503 MW; C7BDB38169D6E3CF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genome polyprotein.
Heparitis C virus.
Viruses; ssNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q81541;
01.NOV-1996 (T.EMBLrel. 01, Created)
01.NOV-1996 (T.EMBLrel. 01, Last sequence update)
01.MAR-2003 (T.EMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100.0%; Score 154; DB 12;
100.0%; Pred. No. 3.7e-14;
artive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR007095; RNLTDD_DS_PS.
InterPro; IPR007095; RNL_PD1_DS_PS.
InterPro; IPR007095; RNL_PD1_DS_PS.
InterPro; IPR007094; RNA_PD1_DS_PS.
InterPro; IPR007094; RNA_PD1_DS_PS.
IPR00709153; HCV_core; 1.
IPR00709153; HCV_NS1; 1.
IPR007091; HCV_NS2; 1.
IPR0070001; HCV_NS3; 1.
IPR007001; HCV_NS4; 1.
IPR00701; HCV_NS4; 1.
IPR00701; HCV_NS5; 1.
IPR00701; HCV_NS5; 1.
IPR00701; HCV_NS4; 1.
IPR00701; HCV_NS5; 1.
IPR00998; VITAL_RGRP; 1.
IPR000009; PP008066; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                               InterPro; IPRO0252; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002531; HCV_core.
InterPro; IPR002531; HCV_MS1.
InterPro; IPR002531; HCV_MS2.
InterPro; IPR004109; HCV_MS3.
InterPro; IPR004109; HCV_MS3.
InterPro; IPR00149; HCV_MS4.
InterPro; IPR00149; HCV_MS4.
InterPro; IPR00166; HCV_MS5.
InterPro; IPR00166; HCV_MS5.
                                                                                                                                                                                                                                                                                                          InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              081541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 7
          RRITH THE RESERVENCE OF STREET OF ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PART BY SERVICE STATE OF THE SERVICE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                    PROSITE; PS00190; CYTOCHROME_C; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
POLYprotein; RNA-directed RNA POLYMETASE; Transmembrane, SEQUENCE 3010 AA; 326793 MW; 3D89304314F9F795 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zheng W.Z.; "Genotype identification of hepatitis c virus (HCV) isolated from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
El and E2/NS1 envelope glycoprotein (Genome polyprotein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zheng W.-Y.Z.;
Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 100.0%; Score 154; DB 12; Best Local Similarity 100.0%; Pred. No. 3.7e-14; Matches 28; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002166; HCV_RdRP.
InterPro; IPR007095; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POl_PSvir.
Pfam; PP01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF01542; HCV_core; Pfam; PF01542; HCV_core; 1. Pfam; PF01569; HCV_core; 1. Pfam; PF01569; HCV_NS2; 1. Pfam; PF02907; HCV_NS3; 1. Pfam; PF01006; HCV_NS4a; 1. Pfam; PF01506; HCV_NS4a; 1. Pfam; PF01506; HCV_NS5a; 1. Pr0Dpcm; PF0198062; VIral_RGRP; 1. Pr0Dcm; PD186062; HCV_NS1; 1. SMART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPRO00745;
InterPro; IPR001490;
InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                 IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
          SEQUENCE FROM N.A STRAIN-HCV-S1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-HCV-N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-HCV-N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                            Lim S.P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q81989
Q81989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 6
```

ô

ô

Gaps

ö

REAL OF STREET O

```
Songsivilal S., Kanistanon D., Kunkitti R.; "Identification and characterisation of Thai isolates of hepatitis {\tt C}
        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                  Length 109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                      109
12318 MW; 9999DBD249BF93B0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1996 (TrEMBLrel. 01, Last Sequence update)
01-CCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                Score 150; DB 12;
Pred. No. 4.1e-15;
1; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 146; DB 12;
Pred. No. 6.5e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 45 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Mismatches
                                                                                                                                                                                                                                                                                                                              28
                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTYRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                97.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           45 45
45 AA; 5014 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94.8%;
                                                                                                                                                                                                                                     Query Match
Best Local Similarity 96.49
Watches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                       109 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-HCV-BB21;
                                                                                                                                                                                      109
                                                                                                                                                                  Polyprotein.
NON_TER 10
SEQUENCE 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein.
NON_TER
SEQUENCE 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 068307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          968308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               068308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                               068307
            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NX OCC DE DATE
                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        <del>a</del>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
Virus Res. 23:39-53(1992).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSIGNS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN CAND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 o polymerase; Transferase; Transmembrane. PUT. P22. PUT. GP35. PUT. GP70. PUT. S22. PUT. NS22. PUT. NS3. PUT. NS3. PUT. NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane
CHAIN 1 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NoV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
(isolate SLI) genomic RNA (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ohno T., Mizokami M.; "Determination of nine genotypes of hepatitis C virus using PCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           05F0B2102CF9DD9D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0%; Score 154; DB 12;
100.0%; Pred. No. 3.7e-14;
iive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     109 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NS4A.
NS4B.
                                                                                                                                                                                                                                                                            InterPro; IPR000745; HCV_NS4a.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR002166; HCV_RGRP.
InterPro; IPR0070166; HCV_RGRP.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_PS-PS.
Pfam; PF01543; HCV_Capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART; SM00487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                           InterPro; IPR002522; HCV_caps1d InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           326564 MW;
                                                                                                                                                                                                                         InterPro; IPR002531; HCV_NS1
InterPro; IPR002518; HCV_NS2
InterPro; IPR004109; HCV_NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                       InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                       InterPro; IPR002519;
                                                                                                                                                                                                                       InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2014 301
3010 AA;
                                                                                                                         P26663; 1JXP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-SL1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     mehtod."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 081807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 8
081807
AC 0818807
DD 01-N0
DD
```

ö

Gaps

; 0

a

```
61
61 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match
Best Local Similarity
Matches 27; Conserv
                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                        NCBI_TaxID-11103;
  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                    STRAIN-RIG256;
                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein.
                                                                                                                                                                                                                                                                                                                     Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QBJYR4
QBJYR4;
                                                                                                                                                                                                                                                       Q8JYS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 14
QBJYR4
ID QBJYR
AC QBJYR
                                                                                                                                                                                                                                         RESULT 13
                                                                                                                                                                                                                                                Q8JYS3
 g
                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                   Gaps
                                                Ø
                                                                                                                                                                                                                                                                                                                         Songsivilal S., Kanistanon D., Kunkitti R.; Identification and characterisation of Thai isolates of hepatitis C
                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
      STRAIN-HCV-BB38;
Songsivilai S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                   Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 45;
                                                                                                                                                  1; Indels
                                                                                                                    CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                             01-NOV-1996 (TFEMBLrel. 01, Last sequence update) 01-OCT-2002 (TFEMBLrel. 22, Last annotation update) Core protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 146; DB 12;
Pred. No. 6.5e-15;
0; Mismatches 1;
                                                                                                                                  Score 146; DB 12;
Pred. No. 6.5e-15;
                                                                                                                                                                                                                        45 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     60 A.A.
                                                                                                                                                  0; Mismatches
                                                                                                                                                                          5 PKPQKKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                 1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                       Created)
                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Genome polyprotein (Fragment).
                                                                                                                                  94.8%;
96.4%;
                                                                                                                   45 AA; 5014 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                              5014 MW;
                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                                               94.88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-OCT-2002 (TrEMBLrel. 22, 01-OCT-2002 (TrEMBLrel. 22, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                       96.48;
                                                                                                                          Ouery Match
Best Local Similarity 96.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best_Local Similarity 96.4 Matches 27; Conservative
                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A. STRAIN-HCV-C94009;
                                                                                                                                                                                                                                                                                                                                                                                                                              45 AA;
                                                                                                                                                                                                                                                                                            NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                       068310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q8JYS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08JYS2
                                                                                                                                                                                                        RESULT 11
Q68310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
QBJYS2
ID QBJYS
AC QBJYS
DT 01-0C
DT 01-0C
DT 01-MA
DE Genom
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
g
                                                                                                                                                                                                                       NA KW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
ö
                                                                                  Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Generic variability of hepatitis C virus in Western Siberia.";
"Generic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPMETEN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MINA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                           Length 60;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 12; Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                   ACCE7D9C8B90299A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B92CCE7D9C8B9029 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                        Score 146; DB 12;
Pred. No. 8.9e-15;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 146; DB 1;
Pred. No. 9e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. 9e-1
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CORE PROTEIN,
                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPORKTKRNTYRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL, AF506612, AAM33380.1, -...
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genome polyprotein (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                                                                                                            Polyprotein.

NON_TER 60 60

SEQUENCE 60 AA; 6725 MW;
                                                                                                                                                                                                                                                                                                                                                                        94.8%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6881 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      94.8%;
                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
```

us-09-491-146a-25.rspt

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kalinina O., Norder H., Mukomolov S., Magnius L.O.;

A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg.;

J. Virol. 76:4034-4043(2002).

- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL: AV701080; AAL58594.1:

InterPro; IPR002522; HCV_capsid.

Profile AP01543; HCV_capsid.
                                                                                                                                                                                          Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (Apr-202) to the EMBL/Geneank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRAN (BY SIMILARITY).

EMBL, AF506621; AAAM33397.1; -
Interpro; IPR002522; HCV_capsid.
Pfan. PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-CT-2002 (TrEMBLrel. 22, Last annotation update)
01-CCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match 94.8%; Score 146; DB 12; Length 7
Best Local Similarity 96.4%; Pred. No. 1.1e-14;
Matches 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                             6EDC082DAE8CB796 CRC64;
01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 146; DB 12;
Pred. No. 1.5e-14;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-21904745; PubMed-11907242;
                                                     Genome polyprotein (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94.8%;
                                                                                                                                                                                                                                                                                                                                                                                                            73 73
73 AA; 8201 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
                                                                                                                                                          SEQUENCE FROM N.A. STRAIN-RIG291;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                         Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
NON_TER
NON_TER
SEQUENCE 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-602;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 15
080P85
AC 080P85;
DT 01-JUN-10
DT 01-JUN-10
DT 01-OCT-10
DE COTE PTC
OC VITUSES
COC NEDACIVO
OC HEPACITAN
OC HEPACITAN
NO NOBLITAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                             FT SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

August 7, 2003, 11:18:58 Search completed: Augus Job time: 26.6364 secs us-09-491-146a-25.rai

```
Appl
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Appl1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 App1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         App]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 25, App
                                                                                                                                                                                                                   (without alignments)
112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                    August 7, 2003, 11:07:41; Search time 10.5455 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      328717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Issued_Patents_AA:*
.: /cgn2_6/ptodata/1/laa/5A_COMB.pep:*
?: /cgn2_6/ptodata/1/laa/5B_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/6A_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/BCTUS_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-635-886C-192
US-08-921-897-23
US-08-290-665A-187
US-08-290-665A-189
US-08-290-665A-190
PCT-US95-10398-187
PCT-US95-10398-187
PCT-US95-10398-187
US-08-88-186-180
US-08-88-186-180
US-08-180-160-2
US-08-180-160-2
US-08-180-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-635-886C-198
US-08-635-886C-232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -US94-07088-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-836-075A-10
                                                                                                                                                                                                                                                                                                                                  154
1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -ns93-08638-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                           protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                              BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                         US-09-491-146A-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query
Match Length DB
                                                                                                                                                                                                                                                                                                         Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sed
                                                                                                                                                                                                                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Minimum DB s
Maximum DB s
                                                                                                                              protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Database
                                                                                                                                                                                 Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ĕ
```

Sequence

```
Sequence 25. Application US/08921887

| Patent No. 6030771
| GENERAL INFORMATION: HUDGANOV YURI E. APPLICANT: FIELDS, HOWAND A. TITLE OF INVENTION: BNDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAN NUMBER OF SEQUENCES: 55
| CORRESPONDENCE ADDRESS: ADDRESS: ADDRESS: ADDRESSE: JONES & ASKEW, LLP STREET: 191 Peachtree Street, N.W., 37th Floor CITY: Allanta
| STREET: GA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                             Sequence 8, Appli
Sequence 8, Appli
Sequence 2, Appli
Sequence 46, Appli
Sequence 10, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 16, Appli
Sequence 16, Appli
Sequence 15, Appli
Sequence 15, Appli
Sequence 15, Appli
Sequence 15, Appli
                                                                                                                                                                                                                                                     Sequence 157,
Sequence 158,
Sequence 159,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OPERATING SYSTEM: PC_DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100.0%; Score 154; DB 3; 100.0%; Pred. No. 1.9e-15;
                                             US-08-904-686A-8
US-08-315-850-8
US-08-501-195-2
US-08-816-075A-46
US-08-816-075A-60
US-08-816-075A-10
US-08-816-075A-15
US-08-816-075A-15
US-08-817-115
US-08-817-115
US-08-817-815
US-08-817-815
US-08-817-815
                                                                                                                                                                                                                                                  US-08-290-665A-157
US-08-290-665A-158
US-08-290-665A-159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STAIL.
COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
AUGITHER: IBM PC COMPALLble
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acids
STRANDEDNESS: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRANDEDNESS: not relevant TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MOLECULE TYPE: protein HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ANTI-SENSE: NO
FRAGMENT TYPE: 1
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; ORGANISM:
US-08-921-887-25
80046444444
80046666666444444
```

å qq

us-09-491-146a-25.rai

```
1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPORKTKRNTYRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   03063-0380
                                                    Sequence 23, Application US/08921887
Patent No. 6030771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 0306:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SS: not relevant not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM: Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          amino acid
                                                                                                                                                                                                                                                                                                                                                                                 USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ANTI-SENSE: NO ORIGINAL SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                          g
                          JS-08-921-887-23
                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           d
                                                                                                                                                                                                           GENERAL INFORMATION:
APPLICANT: LEROUX-ROELS, Geart
APPLICANT: DELES; Robert
APPLICANT: DELES; Robert
TITLE OF INVENTION: VIRUS
TITLE OF INVENTION: VIRUS
TITLE OF INVENTION: VIRUS
FILE REPERENCE: 2752-18
CURRENT APPLICATION NUMBER: US/08/635,886C
CURRENT APPLICATION NUMBER: PCT/EP94/03555
PRIOR FILING DATE: 1996-10-28
PRIOR FILING DATE: 1999-10-48
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PRIOR PLING DATE: 1993-11-04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: DELEYS, Robert
APPLICANT: MARRTENS, Geert
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: UNBUS PROBLES OF HEPATITIS C
TITLE OF INVENTION: UNBUS PROBLES OF HEPATITIS C
TITLE REFERENCE: 2752-18
CURRENT APPLICATION NUMBER: US/08/635,886C
CURRENT FILING DATE: 1996-404-25
PRIOR APPLICATION NUMBER: PCT/EP94/0355
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PALCHIN VERSION 3.1
SEQ ID NO 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 450;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 450;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0%; Score 154; DB 4;
100.0%; Pred. No. 3.3e-14;
11ve 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 100.0%; Score 154; DB 4; Best Local Similarity 100.0%; Pred. No. 3.3e-14; Matches 28; Conservative 0; Mismatches 0;
1 PKPORKTKRNTYRRPODVKFPGGGGIVG 28
                          1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 192, Application US/08635886C Patent No. 655514 GENERAL INFORMATION: APPLICANT: LEROUX-ROBELS, Geert APPLICANT: DELEYS, Robert
                                                                                                                                                                   ; Sequence 191, Application US/08635886C
; Patent No. 6555114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GSGANISM: hepatitis C virus US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 3
US-08-635-886C-192
                                                                                                                                               US-08-635-886C-191
```

å g

g ô

```
GENERAL INFORMATION:
APPLICANT: KHUDVAKOV, YURI E.
APPLICANT: KHUDVAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAM)
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 187, Application US/08290665A
Patent No. 5882852
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 28;
                                                                                                                                                                                                                                                                                                                                                                                    ZIF: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 94.8%; Score 146; DB 3; L. Best Local Similarity 96.4%; Pred. No. 2.7e-14; Matches 27; Conservative 0; Mismatches 1;
```

```
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: WOLLEOTIDE AND DEDUCED
TITLE OF INVENTION: GORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: 200UBNCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
                                                                                                                                                                                                                                                                                                                          DB 2; Length 191;
1.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 2;
                                                                                                                                                                                                                                                                                                                        Query Match 94.8%; Score 146; Best Local Similarity 96.4%; Pred. No. 1 Matches 27; Conservative 0; Mismatche
                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94.8%; Score 146;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/CDCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ыск: US/08/290,665A
15-AUG-1994
N: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: BUKH, J., MILLER, R.H. AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
COMPUTER: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,6658
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 190, Application US/08290665A Patent No. 5882852 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    190:
                               TELEX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 18
SEQUENCE CHARACTERISTICS:
LENGTH: 191 anino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
ORGANIEW
  TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEX: 421792
INFORMATION FOR SEQ ID NO: 19
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: DK12
US-08-290-665A-190
                        (212) 758-48C
(212) 751-6849
                                                                                                                                                                                                                                     ORGANISM: homosaplens;
INDIVIDUAL ISOLATE: S52
US-08-290-665A-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SS: unknown unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NEW YORK
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: unl
ORIGINAL SOURCE
                      TELEPHONE:
TELEFAX: (2
TELEX: 4217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-290-665A-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N.H.

NUCLEOTIDE AND DEDUCED

NUCLEOTE SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 94.8%; Score 146; DB 2; Best Local Similarity 96.4%; Pred. No. 1.9e-13; Matches 27; Conservative 0; Mismatches 1
                                                                                                                                                                                                                            CLASSIFICATION: 435
ATORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEFAX: (212) 756-4800
TELEFAX: (212) 756-4800
TELEFAX: (212) 756-4800
TELEFAX: (212) 756-4800
TELEFAX: (212) 756-6849
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDRESS: unknown
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 PKPQKKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 188, Application US/08290665A
Patent No. 5882852
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
                                                                                                  COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY: USA
2IP: 10154
COMPUTER READABLE FORM:
COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INDIVIDUAL ISOLATE: HK10 US-08-290-665A-187
                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CITY: NEW YORK STATE: NEW YORK
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGINAL SOURCE:
                      USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
US-08-290-665A-188
STATE: N
COUNTRY:
```

ô a

```
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEPATITIS C VIRUS AND THE USE OF REAGENTS DERIVED FROM THESE 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Match 94.8%; Score 146; DB 5; Local Similarity 96.4%; Pred. No. 1.9e-13; les 27; Conservative 0; Mismatches 1.
                                                 R.H.
NUCLEOTIDE AND DEDUCED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DED TITLE OF INVENTION: AMINO ACID SEQUENC TITLE OF INVENTION: AMD THE USE OF REATITLE OF INVENTION: SEQUENCES IN DIAGNA
                              MILLER, R.H. AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: RICHARD W. BORK
REGIESTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 190, Application PC/TUS9510398
GENERAL INFORMATION:
                                                                                                                                                                                                              MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                       IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEX: 421792
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: WORDPERFECT 5.1 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                          345 PARK AVENUE
NEW YORK
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (212) 758-48(
(212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        191 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S52
                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                          ZIP: 10154
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS:
                                                                                                                                                                   NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES:
                                                                TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                unknown
                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: amino acid
STRANDEDNESS: un)
        GENERAL INFORMATION:
APPLICANT: BUKH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: unk
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PCT-US95-10398-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PCT-US95-10398-190
                                                                                                                                                                                                              ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                          COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX:
                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                         ö
                         Gaps
                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: NUCLECTIDE AND DEDUCED

TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             94.8%; Score 146; DB 5; Length 191;
96.4%; Pred. No. 1.9e-13;
tive 0; Mismatches 1; Indels
  Pred. No. 1.9e-13;
0; Mismatches 1; Indels
                                                              1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                   PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDIUM TYPE: FLORPY DISK
COMPUTER: IBM PC CONDAILELE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-ANG-1995
CLASSIFICATION:
PROR APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PROR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W: BORK
REGISTRATION NUMBER: 36,436
                                                                                                                                                                                                                                                APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 9
PCT-0S95-10398-188
; Sequence 188, Application PC/TUS9510398
                                                                                                                                                                                                       Sequence 187, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                SSEE: MORGAN & FINNEGAN
7: 345 PARK AVENUE
NEW YORK
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEX: 421792
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEFHONE: (212) 758-4800
TELEFAX: (212) 751-6849
  96.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: HK10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
Best Local Similarity 96.49
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 96.4 Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                             CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
                                                                                                                                                                                    -US95-10398-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PCT-US95-10398-187
                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE:
STREET: 34
                                                                                                                                                                                                                                                                     APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CITY:
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                 RESULT
                                                                                                  a
                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

```
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: LERONY-ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: DELEYS, Robert
TITLE OF INVENTION: URMUNDCMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VIRUS
FILE REFERENCE: 275-18
CURRENT FILING DATE: 1996-04-25
PRIOR APPLICATION NUMBER: DET/EP94/03555
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PATENTIN VERSION 3.1
SEQ ID NO 230
LENGTH: 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; DB 3;
3.2e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; Score 146; DB
; Pred. No. 3.2e
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME/KEY: MISC_FEATURE
LOCATION: (156)..(157)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LOCATION: (144)..(144)
OTHER INFORMATION: Xaa is any amino acid
                                                             FILING DATE: 23 Oct 1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
NAME: KAMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REGISTRATION NUMBER: 29,775
INFORMATION FOR SEQ ID NO: 44:
FILING DATE: 21 Apr 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/04155
FILING DATE: 23 Oct 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-635-886C-230
; Sequence 230, Application US/08635886C
; Patent No. 6555114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 94.8%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                   i: 319 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY: MISC_FEATURE
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is
                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                       ; TOPOLOGY: linear;
; MOLECULE TYPE: peptide
US-08-836-075A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEY: MISC_FEATURE LOCATION: (167)..(167)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 44, Application US/08836075A
Patent No. 6180768
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: STUTYER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: AGENTS
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 146; DB 5; Length 191;
Pred. No. 1.9e-13;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
Microsoft Word 6.0 / ASCII text output
                                                                                                     COMPUTER READBLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: HONDPERFECT 5.1
CURRENT PAPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTONNEY/AGBNT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REGISTRATION NUMBER: 36,459
REGISTRATION NUMBER: 36,459
RECISTRATION NUMBER: 36,459
TELECOMUNICATION INFORMATION:
TELECOMUNICATION INFORMATION:
TELECOMUNICATION INFORMATION:
TELECOMUNICATION INFORMATION:
TELECOMUNICATION INFORMATION:
TELECOMUNICATION OF SEQ ID NO: 190:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
STRENDEDNESS: UNKNOWN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,075A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: homosapiens
;
INDIVIDUAL ISOLATE: DK12
PCT-US95-10398-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Floppy disk
    STREET: 345 PARK AVENUE CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: unknown ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OPERATING SYSTEM:
SOFTWARE: Microso
                                                 NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CITY: HOUSTON STATE: TEXAS
                                                                       USA
                                                                                          10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-836-075A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER:
```

à g

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                         Score 145; DB 3; Length 34;
Pred. No. 4.5e-14;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,846
FILING DATE: 09-FEB-1998
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 9-027015
FILING DATE: 10-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-024045
FILING DATE: 09-FEB-1996
ATTORNEX/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 145; DB 4; L
Pred. No. 5.7e-14;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: YAMAGUCHI, Kenjiro
APPLICANT: KASHIMAKUMA, Tomiko
APPLICANT: CHIBA, Yukie
APPLICANT: CHIBA, Yukie
APPLICANT: YAGI, Shintaro
APPLICANT: HASEGAMA, Akira
TITLE OF INVENTION: CHIMERA ANTIGEN PEPTIDE
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                              4 PKPQKKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: Wegner, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 053466/0225
TELECOMMUNICATION:
                                                                                                      ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Human Hepatitis C Virus
US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; Sequence 36, Application US/09020846; Patent No. 6322965; General INFORMATION; APPLICANT: YAMAGUCHI, Keniiro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3000 K Street, N.W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FOLEY & LARDNER
                                                                                                                                                                                                                                     94.2%;
1larity 96.4%;
Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 94.2%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (202) 672-5399
                           uss: single
linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     43 amino acids
                                               TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY: U.S.A.
21P: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO
                 amino acid
                                                                                                                                                                                                                                Ouery Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITY: Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid STRANDEDNESS:
            TYPE: amino a STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: DALBON, PASCAL
APPLICANT: JOLIVET, MICCHI
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY
TITLE OF INVENTION: FOR DETECTING THE LATTER
CORRESPONDENCE: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94.8%; Score 146; DB 4; Length 319; 96.4%; Pred. No. 3.2e-13; 1ve 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQKKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                              FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (174)..(174)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                            FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (232)...(232)
OTHER INFORMATION: Xaa is any amino acid
INFORMATION: Xaa is any amino acid
                                                           LOCATION: (171)..(172)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                NAME/KEY: MISC_FEATURE LOCATION: (177)...(177) OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: MISC_FEATURE
LOCATION: (233)..(233)
OTHER INFORMATION: Xaa is any amino acid
08-635-886C-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION: 530
PRIOR APPLICATION DATE:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
REGISTRATION NUMBER: 30,024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER: WPB 28682
TELECOMOUNICATION INFORMATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : Sequence 6, Application US/08380160
; Patent No. 6235284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. Box 19928
CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 96.4*
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                        NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STATE: VI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 13
```

ò q

```
Search completed: August 7, 2003, 11:23:50
Job time : 10.6364 secs
                                                                                                                                                                                                                          Michel
SYNTHETIC POLYPEPTIDES BELONGING TO THE
SYNTHETIC C VINUS (HCV) AND WHICH CAN BE USED ESPECIALLY
FOR DETECTING THE LATTER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94.2%; Score 145; DB 3; Length 44; 96.4%; Pred. No. 5.9e-14; ive 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:

WEDDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATIOG SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
FILING DATE:
CLASSIFICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATONNEY/AGENT INFORMATION:
REGISTRATION NUMBER: 30,024
RELEPHONE: (703)836-6400
TELEFRAX: (703)836-6400
1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: Human Hepatitis C Virus STRAIN: H77
                                                                                                                                      Sequence 1, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
APPLICANT: DALBON, Pascal
TITLE OF INVENTION: SYNTHERIC POLY
TITLE OF INVENTION: HEPATITIS C VI
TITLE OF INVENTION: CON DETECTING
NUMBER OF SEQUENCES: 12
                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. Box 19928
CITY: Alexandila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FRAGMENT TYPE: N-terminal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                                                                              Κ,
                                                                                                                                                                                                                                                                                                                                                                                                            STATE: VA
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LOCATION:
    ò
                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
```

g

Run on:

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries length: 0 length: 2000000000 Minimum DB seq Maximum DB seq

451899

tal number of hits satisfying chosen parameters:

451899 seqs, 118759770 residues

Searched:

'(gn2\_6/ptodata//Pubpaa/US07\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Pubpaa/US07\_NEW\_PUB.ppp:\*
'(gn2\_6/ptodata//Pubpaa/US06\_NEW\_PUB.ppp:\*
'(gn2\_6/ptodata//Pubpaa/US06\_NEW\_PUB.ppp:\*
'(gn2\_6/ptodata//Pubpaa/US07\_NEW\_PUB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US08\_NEW\_PUB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US08\_NEW\_PUB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US08\_NEW\_PUB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US08\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US08\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US09\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US09\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US09\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US09\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US09\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US10A\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US10A\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US10A\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US10A\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US10A\_PUBCOMB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US10A\_PUB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US10A\_PUB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US10A\_PUB.ppp:\*
'(gn2\_6/ptodata//Tpubpaa/US10A\_PUB.ppp:\* Published\_Applications\_AA:\* Database :

ALIGNMENTS

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Sequence 1, Applisequence 1, Applisequence 10, Applisequence 7, Applisequence 77, Applisequence 14, Applisequence 6, Applisequence 6, Applisequence 15, Applisequence 152, Applisequence 153, Applisequence 154, Applisequence 154, Applisequence 155, Applisequence Sequence 44, Appl 42, Sequence , Description 0 US-09-851-138-44 5 US-10-367-677-1 0 US-09-758-308-1 US-09-758-308-1 0 US-09-758-875-8 1 US-09-921-397-77 US-09-921-397-77 US-09-921-397-78 US-09-951-138-46 US-09-951-138-46 US-09-951-138-46 US-09-981-138-46 US-09-981-152 US-09-981-152 US-09-899-046-152 US-09-899-046-152 US-09-899-046-152 US-09-899-046-152 US-09-899-046-142 00000011111 Length DB Query Match I Score Result Š

ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ 1D NO: 44:

Appl Appl

Sequence 44, Appl Sequence 2, Appli Sequence 2, Appli Sequence 3, Appli Sequence 76, Appli	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	22000000000000000000000000000000000000	20, 4 , 4 , 4 , 4 , 4 , 4 , 4 , 4 , 4 , 4
US-09-878-281-44 US-09-929-955-2 US-10-104-966-2 US-09-194-949-3 US-09-851-138-76	09-851-13 09-851-13 09-899-04 09-899-04 09-899-04	-09-878-281 -09-878-281 -09-878-281 -09-873-025 -09-899-303 -09-995-860 -09-995-860	US-10-259-275-40 US-09-742-659-4 US-09-925-572-9 US-09-929-955-1 US-09-747-419-20 US-09-891-894-3 US-10-104-966-1 US-10-259-275-20 US-09-238-076-2
111111	20111111	111101111	
169 182 182 191	3310 3110 3110 3110 3110 3110	3119 3119 3119 8009 8009 2809	2985 3011 3011 3011 3011 3011 3011
99 99 99 99 99 99 99 99 99 99 99 99 99	99999999999999999999999999999999999999	00000000000000000000000000000000000000	4444444444 2.2.2.2.2.2.2.2.2.2.2.2.2.2.2
145 145 145 145	1145 1445 1445 1455 1455	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	11111111111111111111111111111111111111
16 17 18 20	222 222 224 24	9 9 9 9 9 9 9 8 9 9 9 9 9 9 9 9 9 9 9 9	3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4

## Sequence 44, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION: APPLICANT: MAENTENS, GEERT STUTVER, LIEVEN TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROFHYLACTIC, THERAPEUTIC AND DIAGN ADDRESSES ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433 CITY: HOUSTON STATE: TEXAS COUNTRY: USA ZIP: 77210-443 COMPUTER READBLE FORM: MEDIUM TYPE: Floppy disk COMPUTER: READBLE FORM: MEDIUM TYPE: Floppy disk CURRENT APPLICATION DATA: APPLICATION UNDRER: US/09/851,138 FILING DATE: 09-MBy 2001 PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/836,075 FILING DATE: <Unknown> APPLICATION NUMBER: EP 94870166.9 FILING DATE: 21 Oct 1994 APPLICATION NUMBER: EP 95870076.7 FILING DATE: 28 Jun 1995 AGENTS NUMBER OF SEQUENCES: 207 CORRESPONDENCE ADDRESS US-09-851-138-44 RESULT 1

us-09-491-146a-25.rapb

```
GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C \( \)
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 14114.034912
CURRENT APPLICATION NUMBER: US/09/758,308
CURRENT APPLICATION NUMBER: 60/092,339
PRIOR APPLICATION NUMBER: 60/092,339
PRIOR FILING DATE: 1999-07-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PATENTI NOS: 5
SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 94.2%; Score 145; DB 10; Length 74; Best Local Similarity 96.4%; Pred. No. 1.7e-13; Matches 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                         COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 145; DB 9;
Pred. No. 2.2e-13;
                                                                                                                                                                                                                                                                                      CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: <21 Oct 1994
APPLICATION NUMBER: EP 97870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
                                                 WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775
                                                                                                                            COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 1, Application US/09758308 Patent No. US20020090607A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 74 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
0
                                                                   STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                 ADDRESSEE: ARNOLD,
  NUMBER OF SEQUENCES: 20 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: PRT ORGANISM: Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: MAERTENS, GEERT STUYVER, LIEVEN TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILE REFERENCE: 103959
CURRENT APPLICATION NUMBER: US/09/389,756
PRIOR APPLICATION NUMBER: US/09/389,756
PRIOR PELING DATE: 1999-09-07.
PRIOR PELING DATE: EARLIER FILING DATE: 1999-03-05
NUMBER OF SEQ ID NOS: 11
SEQ ID NO: 1
SEQ ID NO: 1
LUNCH: 44
                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
0
                                                                                                                                                                               Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 44;
                                                                                                                                                                                                                            Indels
                                                                                                                                                                        Query Match 94.8%; Score 146; DB 10; Best Local Similarity 96.4%; Pred. No. 6.2e-13; Matches 27; Conservative 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 145; DB 15;
Pred. No. 9.9e-14;
0; Mismatches 1;
  LENGTH: 319 amino acids
TYPE: amino acid
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 44:
                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                           5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 10, Application US/09851138 Publication No. US20020183508A1
                                                                                                                                                                                                                                                                                                                                                                                                                          : Sequence 1, Application US/10367677 iblication No. US20030118604A1 ENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AGENTS
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94.2%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: Hepatitis C virus PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAGES: 3392-3396
DATE: 1991
; RELEVANT RESIDUES: 2 TO 45
US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 96.41
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: JOLIVET, MICHEL APPLICANT: PENIN, FRANCOIS APPLICANT: DALBON, PASCAL APPLICANT: LADAVIERE, LAUI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GENERAL INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                               RESULT 2
US-10-367-677-1
                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

Gaps

ö

```
APPLICANT: MAERTENS, GEERT STUVER, LIEVEN STUVYER, LIEVEN TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE: TEXAS

CONTRY: USA
2 IP: 77210-43

ZONDUTRY: USA
COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION DATA:
RELIGATION UNBER: US/09/851,138

FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                   Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 108;
                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                 Score 145; DB 10;
Pred. No. 2.5e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94.2%; Score 145; DB 10; 96.4%; Pred. No. 2.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
PAPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
INFORMATION FOR SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                  1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-851-138-14; Application US/09851138; Sequence 14, Application US/09851138; Publication No. US20020183508A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 78, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHIT
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGIH: 108 amino acids
                                                                 Ouery Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS
; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 96.49
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 8
US-09-921-397-78
                                                                                                                                                          ò
                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
TITLE OF INVENTION:
TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
TITLE OF INVENTION: applications thereof
FILE REFERENCE: B4609A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2000-08-03
REIOR APPLICATION NUMBER: EP 00402225.7
REIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 97;
                                                                   Sequence 8, Application US/09756875
Patent No. US20020150990A1
GENERAL INFORMATION:
APPLICANT: PIKE, IAN
IIILE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 701-E Columbia Square
STRRET: 555 13th Street, N. W.
STRATE: D. C.
                                                                                                                                                                                                                                                                                                                                                                                                                                        OPERATING STATEM: OFFICE/AS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 145; DB 10;
Pred. No. 2.3e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-AUG-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTORNEY.AGENT INFORMATION:
NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1808-157A
TELECOMMUNICATION:
TELECOMMUNICATION:
TELECHHONE: (202)/83-6040
INFORMATION FOR SEQ ID NO: 8:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; Sequence 77, Application US/09921397
; Patent No. US20020151484A1
                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: unknown

MOLECULE TYPE: peptide
US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                           D. C.
                                                                                                                                                                                                                                                                                                                              COUNTRY: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQ ID NO 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
```

ö

```
STUTVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPOTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER: IEMP C COMPATIBLE
COMPUTER: IEM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 6.0 / ASCII text output
CURRUT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: B 94870166.9
FILING DATE: CUNKNOWN>
APPLICATION NUMBER: B 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: E 95870076.7
FILING DATE: 22 Jun 1995
ATTORNEY/AGENT INFORMATION:
NAMBER: RAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REGISTRATION NUMBER: 29,775
REFERENCE/POCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 138;
                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Indels
                  ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 145; DB 10;
Pred. No. 3.4e-13;
                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
SIREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Mismatches
                                                       1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                              1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY: Ilnear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; Sequence 152, Application US/09899046; Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                           Sequence 60, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 138 amino acids
                ö
                                                                                                                                                                                                                                                                           APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              94.28;
96.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match
Best Local Similarity 96.4
Matches 27; Conservative
           27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-899-046-152
                                                                                                                                                               RESULT 10
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-851-138-60
             Matches
                                                         ò
                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: MAGALIN, CONTROL STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                          ö
         polypeptides selected
the hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPURE: TEAAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER: USA
EDIDW TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICTOSOFT WORD 6.0 / ASCII text output
SOFTWARE: MICTOSOFT WORD 6.0 / ASCII text output
APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 2 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY AGENT INFORMATION:
NAME: KAMMERER, PATRCIA A.
REGISTRATION NUMBER: 29,775
REGISTRATION NUMBER: 29,775
REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                 Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 137;
                                                                                                                                                                                                                                                                                                                                                                                                          1; Indels
                                                                                                                                                                                                                                                                                                                                                          Score 145; DB 10;
Pred. No. 2.8e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 145; DB 10;
Pred. No. 3.4e-13;
TITLE OF INVENTION: SID nucleic acids and polifice of INVENTION: pathogenic strain of the TITLE OF INVENTION: pathogenic strain of the TITLE OF INVENTION: applications thereof the TITLE REFERENCE: B4609A - JAZ CURRENT APPLICATION NUMBER: US/09/921,397; CURRENT FILING DATE: 2001-08-02; PRIOR FILING DATE: 2000-08-03; NUMBER OF SEQ ID NOS: 156; SOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-851-138-46; Sequence 46, Application US/09851138; Publication No. US20020183508A1; CENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH: 137 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94.28;
96.48;
                                                                                                                                                                                                                                                                                                                                                          Query Match

Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                       ; ORGANISM: Hepatitis C virus US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                  LENGTH: 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-851-138-46
                                                                                                                                                                                                                           SEQ ID NO 78
                                                                                                                                                                                                                                                                       TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3
```

Gaps

ö

g

```
New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                                                         ; DB 11; Length 169; 4.3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
    MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OPERATING SYSTEM: PC-DOS/MS-DOS
SOFWWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94.2%; Score 145; DB 11; 96.4%; Pred. No. 4.3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 94.2%; Score 145; DE
Best Local Similarity 96.4%; Pred. No. 4.3e
Matches 27; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
FILING DATE:
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 44, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 42, Application US/09878281
Publication No. US20030032005A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: New sequence TITLE OF INVENTION: GENOLOGYPES COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk COMPUTER: IBM PC COMPATILLE COMPUTER: IBM PC COMPATILLE OPERATING SYSTEM: PC-DOS/MS-DC
                                                                                                                                                                                                                             TYPE: amino acids TOPOLOGY: 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : 169 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 96.4 Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                  ) MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: 91
TITLE OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                           Score 145; DB 11; Length 166;
Pred. No. 4.2e-13;
0; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PC-DOS/MS-DOS
SOFTWARE: PELLATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match 94.2%; Score 145; DB 11; Best Local Similarity 96.4%; Pred. No. 4.2e-13; Matches 27; Conservative 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPORKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: US/09/878,281 FILING DATE: PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 152, Application US/09878281 Publication No. US20030032005A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 42, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION: APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acids
                                                                                                                                                                                                INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-878-281-152
                                                                                                                                                                                                                                                                                                           ) MOLECULE TYPE: protein US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQUENCES: 2
COMPUTER READABLE FORM:
                                                                                                                                                                                FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 13
US-09-899-046-42
```

Gaps

ö

ó 셤

```
us-09-491-146a-25.rapb
```

```
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: IBM C compatible
COMPUTER: IBM C compatible
CONFURENT IBM C compatible
CONFURENT PAPELICATION DATA:
CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
FILING DATE:
```

polyprotein polyprotein

genome

polyprotein polyprotein

genome genome genome

genome genome genome genome genome genome genome

```
Use Core to the protein core protein cylonoly. The core protein cylonoly yalvanety to genotype 2, N2

A; Abriety; genotype 2, N2

C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000

C; Accession: 841353

R; Wan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.

Submitted to the EMBL Data Library, January 1994

A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.

A; Reference number: $4131

A; Accession: 841353

A; Accession: 841353

A; Residues: 1-108 < VAN.

A; Residues: EMBL: 229456

A; Experimental source: genotype 2, N2

C; Superfamily: hepatitis C virus genome polyprotein

C; Keywords: capsid protein; core protein; polyprotein

F; 1-108/Product: core protein #status predicted <NAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                M.Contains: core protein
C.Species: hepatitis C virus (genotype 2, N4) (fragment)
C.Species: hepatitis C virus
A.Variety: genotype 2, N4
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C.Accession: S41355
R.Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A.Reference number: S41355
A.Accession: S41355
                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus (genotype 2, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 142; DB 2; Length 108;
Pred. No. 1.4e-12;
); Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              91.0%; Score 142; DB 2; Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A Molecule type: genomic RNA
A; Residues: 1-108 < VAN>
A; Cross-references: BMBL:229458
A; Experimental source: genotype 2, N4
C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F;1-108/Product: core protein *status predicted < VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Mismatches
                                                                                                                                                                                                                                                                                                                                   ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
PC2219
GNWVTC
GNWVCJ
S18030
                                                                      GNWVCH
S40770
JC5620
JO1303
GNWVJ8
S41351
S41349
S41352
JOHNWYCS
S21336
S41367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 92.9
Matches 26; Conservative
 876
3010
3010
3010
3010
3011
3011
115
115
115
640
991.0
991.0
991.0
991.0
991.0
991.0
991.0
991.0
991.0
991.0
991.0
991.0
 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 2
S41355
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein
polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein - hepa
polyprotein - hepa
genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein
                                                                                                                       August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome
genome
genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome
genome
genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         депоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qenome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome
                                                                                                                                                                                                                                                                                                                                                                 283308
               GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                             tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                  156
1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                           283308 segs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                         protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       $841353
$8413353
$8413355
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$841346
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
$84136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JH0711
S19876
                                                                                                                                                                                                                                                                        BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     518032
                                                                                                                                                                                                                                                                                                                                                                                                   seq length: 0 seq length: 2000000000
                                                                                                                                                                                                 US-09-491-146A-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR_76:*
1: pirl:*
2: pir2:*
3: pir3:*
4: pir4:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Perfect score:
                                                                                                                                                                                                                                                                        Scoring table:
                                                                                         protein
                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB
Maximum DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Database :
                                                                                                                                                                                                                                     Sequence:
                                                                                                                                                                                                                                                                                                                           Searched:
                                                                                                                           Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
```

ö

Gabs

2

ö

```
RESULT
S41370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                         ö
                                                                                                                                                                                                                 Genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: $41357
Submitted to the EMBL Data Library, January 1994
A;Bescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: $41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41348
R;van Doorn, L.J.; Rleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
`Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
'eference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                         Gaps
                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 2; Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                         Indels
Pred. No. 1.4e-12;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      lecule type: genomic RNA
A;Cross-references: BNBL:229460
A;Experimental source: genotype 2, N6
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Molecule type: genomic RNA
A;Residues: 1-108 <VAN>
A;Cross-references: EMBL:229451
A;Experimental source: genotype 1, N6
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      , Match 91.0%; Score 142; DB 2; I
Local Similarity 92.9%; Pred. No. 1.4e-12;
les 26; Conservative 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match 91.0%; Score 142; DB 2; I
Best Local Similarity 92.9%; Pred. No. 1.4e-12;
Matches 26; Conservative 0; Mismatches 2;
                                                                 1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                         1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PKPORKPNRNTNRRPODVKFPGGGOIVG
92.98;
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Accession: S41348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                 ò
                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
N;Contains: core protein
C;Species: hepatitis, C virus
A;Variety: genotype 5, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (genotype 1, N1) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N1
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
A;Variety: genotype 5, N5
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Accession: S41341
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Description: S41341
A; Accession: S41341
A; Accession: S41341
A; Recience number: S41341
A; Recience: EMBL: Z29444; NID: g443850; PIDN: CAA82582.1; PID: g443851
A; Experimental source: genotypel, NI
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-112/Product: core protein *status predicted <AMT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: genomic RNA
A; Residues: 1-114 <VAN>
A; Cross-references: EMBL: Z29473; NID: g443908; PIDN: CAA82611.1; PID: g443909
A; Experimental source: genotype 5, N4
C; Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                        LiPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Lipa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 by
                                                                                                                                                        þ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 112;
                                                                                                                                                        to 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C; Accession: S41370
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to
A; Reference number: S41341
A; Accession; S41370
                                                                                 Syncogonia, 1.1.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A:Description: Analysis of hepatitis C virus genotypes 1
A:Reference number: S41341
A:Accession: S41371
A:Accession: S41371
A:Residues: I-112 < VANA
A:Cross-references: EMBL:Z29474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 142; DB 2,
No. 1.5e-12;
                                                                                                                                                                                                                                                                                                                           A.Experimental source: genotype 5, N5
C.Superfamily: hepatitis C virus genome polyprotein
C.Keywords: capsid protein; core protein; polyprotein
F:1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 142; DB 2;
Pred. No. 1.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match 91.0%; Score 142; DE Best Local Similarity 92.9%; Pred. No. 1.5e Matches 26; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
```

#text\_change 17-Nov-2000

LIPA.

5 by

```
N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: $4134 #teter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Reference number: $4134
A; Reference number: $41344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A)Cross references: EMBL:229445; NID:9443852; PIDN:CAA82583.1; PID:9443853
A)Experimental source: genotype 1, N10
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Keywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein #status predicted <NAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Cross-references: EMBL:Z29447; NID:g443856; PIDN:CAA82585.1; PID:g443857
                                 polyprotein - hepatitis C virus (genotype 1, N10) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein – hepatitis C virus (genotype 1, N2) (fragment)
                                                       N.contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N10
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_chan
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_chan
C;Accession: S41342
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 3
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: genomic RNA A; Residues: 1-115 < VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                  A; Accession: S41342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 12
                                            genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Genome polyprotein - hepatitis C virus (genotype 5, N2) (fragment)
N.Contains: core protein
C.Speciaes: hepatitis C virus
A.Variety: genotype 5, N2
Ate: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C.Cessaion: 841368
N.Yan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A.Pescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A.Reference number: 841368
A.Accession: 841368
A.Accession: 841368
A.Accession: 941368
A.Accessio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (genotype 5, N3) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Pariety: genotype 5, N3
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
`.Accession: $41369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           January January 1994

Appearithion: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.

A;Reference number: S41341

A;Reference number: S41341

A;Reference number: S41341

A;Reference number: S41340

A;Reference number: S41340

A;Residues: 1-114 </Pre>

A;Residues: 1-114 

A;Residues: 1-114 

A;Residues: 1-114 

C;Superfamily: hepatitis C virus genome polyprotein

C;Reywords: capsid protein; core protein; polyprotein

E;1-114/Product: core protein #status predicted 

A;Experimental source: genotype 5, N3

C;Reywords: capsid protein; status predicted 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 114;
                                                                                                        Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 114;
                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                Score 142; DB 2; Lu
Pred. No. 1.5e-12;
0; Mismatches 2;
C; Keywords: capsid protein; core protein; polyprotein F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
91.0%; Score 142; DB 2;
Best Local Similarity 92.9%; Pred. No. 1.5e-12;
Matches 26; Conservative 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 142; DB 2;
Pred. No. 1.5e-12;
0; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                          1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                           91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 91.0
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                         1 Similarity 92.9
26; Conservative
                                                                                                           Query Match
                                                                                                                                             Best Local
                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
S41369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

ö

Gaps

ö

Score 142; DB 2; Length 115; Pred. No. 1.5e-12; 0; Mismatches 2; Indels

91.0%; 92.9%;

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     #text_change 17-Nov-2000
                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LIPA
                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein - hepatitis C virus (genotype 1, N8) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Accession: $41350
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by A; Reference number: $41341
A; Accession: $41350
                                                                                                                              Length 115;
                                                                                                                                                                            Indels
                                                                                                                         Score 142; DB 2; I
Pred. No. 1.5e-12;
); Mismatches 2;
A; Experimental source: genotype 1, N2 C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F; 1-115/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (genotype 1, N;Contains: core protein S; Species: hepatitis C virus A;Variety: genotype 1, N8 C;Date: 19-May-1994 #sequence_revision 26-Jul-1995
                                                                                                                                                                                                                                                             PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                               1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                            ö
                                                                                                                           91.0%;
ilarity 92.9%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: genomic RNA
```

10

RESULT

```
N; Contains core protein
C; Species: hepatitis C virus
A; Variety: genotyte 1, N5
C; Date: 19-May 1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                         A; Reference number: S41341
A, Accession: S41347
A; Molecule type: genomic RNA
A; Residues: 1.115 < VANA
A; Residues: 1.115 < VANA
A; Cross-references: EMBL: Z29450; NID: 9443862; PIDN: CAA82588.1; PID: 9443863
A; Experimental source: genomy poly NS
G; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted < WAT>
                                                                                                                                                                                                                                                                                                               R) van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994 A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by Lipa.
                                                                                                                                                   genome polyprotein - hepatitis C virus (genotype 1, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 142; DB 2;
Pred. No. 1.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Mismatches
                   PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Search completed: August 7, 2003, 11:21:47
Job time : 9.54545 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                         C; Accession: S41347
                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                             GA1354

genome polyprotein - hepatitis C virus (genotype 2, N3) (fragment)
Nicontains: core protein
Netles: hepatitis C virus
ridety: genotype 2, N3
Liety: genotype 3, N3
Liety: genotype 2, N3
Liety: genotype 3, N3
Rivan Doorn, LiJ.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by Liph.
A.Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Species: hepatitis C virus
A;Variety: genotype 1, N3
C;Date: 19-May 1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
A;Residues: 1-115 <VAN>
A;Cross-references: EMBL:Z29453; NID:g443868; PIDN:CAA82591.1; PID:g443869
A;Experimental source: genotype 1, N8
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Cross-references: EMBL:229448; NID:g443858; PIDN:CAA82586.1; PID:g443859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C; Accession: $41345
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by Lipp.,
A; Reference number: $41341
                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           nome polyprotein - hepatitis C virus (genotype 1, N3) (fragment)
                                                                                                                                                                                  Score 142; DB 2; Length 115;
Pred. No. 1.5e-12;
0; Mismatches 2: Todal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 142; DB 2; Length 11.
Pred. No. 1.5e-12;
Timmatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 142; DB 2; Length 11
Pred. No. 1.5e-12;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Experimental source: genotype 2, N3
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A:Experimental source: genotype 1, N3_
C;Superfamily: hepatitis C virus genome polyprotein
(Expevords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Mismatches
                                                                                                                                                                                                                                                                                                                                      1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   91.0%;
92.9%;
                                                                                                                                                                                                   91.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Cross-references: EMBL: 229457
                                                                                                                                                                                                                       Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity 92.9
les 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: genomic RNA A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: S41354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Accession: S41345
                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ontains:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
```

Gaps

ö

1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28

ö

Gaps

ö

Length 115; Indels

```
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec Run on:

US-09-491-146A-26 156 1 PKPORKPURNTURRPQDVKFPGGGQIVG 28

Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched:

stal number of hits satisfying chosen parameters:

127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

		ption	hepatitis c	h genome po		h genome po	h genome po	h genome po	h genome po	c	ď	0	ㄷ	homo sapien		archaeoglob	bos taurus				plasmodium	drosophila	mus musculu		shallot vir	rattus norv	mns mnscnjn	rattus norv	_		rattus norv				
		Description	P27959	001404	001403	P27960	P27961	P26663	P26662	P27958	P26660	P26661	P26664	P29846	000269	028653	014738	P51816	P45978	028715	P55106	P21573	080361	P06915	P23093	P49415	P97820	P14771	004572	P58405	P47708		Q9erg2	P25220	Q9jmc1
3		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1															_	_			_														
OT THE PLANT		ID	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJS	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCV1	POLG_HCVTW	POLG_HCVJT	2A5D_RABIT	2A5D_HUMAN	FMR2_HUMAN	SCD6_YEAST	YF57_ARCFU	GDF6_BOVIN	YB1_XENLA	RK4_TOBAC	CSP_PLABE	CSP_PLABA	SDC_DROME	M4K4_MOUSE	SC25_YEAST	COAT_SHVX	STR3_RAT	RP3A_MOUSE	RP3A_RAT	- 1	POLS_IBDVP	P5PA_RAT
		8	-		٦		-	-	7	٦	П	٦	~	7	٦	П	7	-	-	-	-	П	٦	-	-	7	7	-	-1	-	~	-	-4	-	Н
		Match Length	513	520	520	737	737	3010	3010	3011	3033	3033	3011	3010	3010	586	602	1311	349	167	436	303	282	339	347	395	1233	1252	262	461	681	684	196	993	1001
œ	Query	Match	91.0	91.0				91.0	91.0	91.0	91.0					35.6	34.3	34.0	33,3	32.7	32.7	32.4	32.1	32.1	32.1	31.4	31.4	31.4	30.8					•	30.8
		Score	142	142	142	142	142	142	142	142	142	142	138	135	134	55.5	53.5	53			ın	50.5	20	50	20	49	49	49	48	48	48	48	48	4.8	8 7
	Result	Q	П	7	m	4	ß	9	7	80	σ.	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	32	33

				Q49378 mycoplasma							
POLS_IBDV5	POLS_IBDVC	POLS_IBDVS	C3G_DROME	P32_MYCGA	XG_HUMAN	RNF4_HUMAN	RS7_FUGRU	GRAB_MOUSE	LIPA_CAUCR	IE63_HSV2H	SYN_BRUMA
-	-	-	н	m	7	٦		٦	٦	-	П
1012	1012	1012	1571	304	180	190	194	247	325	512	548
œ	8.0	8.0	30.8	30.4	30.1	30.1	30.1	30.1	30.1	30.1	30.1
30.	m	""									
48 30.1	48 3		48	47.5	47	47	47	47	47	47	47

## ALIGNMENTS

HCV02_HCV02 STANDARD; PRT; 513 AA. POLC_HCV02 STANDARD; PRT; 513 AA. POLC_HCV02 STANDARD; PRT; 513 AA. 01-AUG-1992 (Rel. 23, Created) 01-AUG-1992 (Rel. 23, Last sequence update) 16-OCT-2001 (Rel. 40, Last annotation update) 16-OCT-2001 (Rel. 40, Last annotation update) Envelope glycoprotein [Contains: Capsid protein C (Core protein) (P22); Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1) (Fragment). Hepatitis C virus (isolate HC-02) (HCV). Hepatitiss; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;	Yamamoto K., Lizuka H., Labiro S.; Lis C virus genome having poor mparative study of four distinct mS2A, NS2B, NS4A AND NS4B ARE SIBLE MEMBRANE-RELATED FUNCTION. N THE VIRAL RNA REDLICATION. IRUS IS A NUCLECCAPSID COVERED BY A ELOPE CONSISTS OF TWO PROTEINS: THE NUCLECCAPSID IS A COMPLEX OF THE NUCLECCAPSID IS A COMPLEX OF THE TAS A COMPLEX OF	use by non-prolit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  EMBL; Dl0074; BAA00968.1; InterPro; IPR002521; HCV_core. InterPro; IPR002521; HCV_core. InterPro; IPR002521; HCV_core. InterPro; IPR002519; HCV_c
STANDARD: (Rel. 23, Cre (Rel. 23, Las (Rel. 40, Las roctein [Conte coprotein El (NSI)] (Fre virus (Isolative es (INA positive es	TAXID=1111;  ENCE FROM N.A.  INE=92230232; PubMed=1314459; OCO H., Kurai K., Okada K.I., Ata T., Fukuda S.T., Tadda F., M.  1-length sequence of a hepatilody last to reported isolates: co types."; I-length sequence of a hepatilody 188:331-341(1992).  FUNCTION: THE SMALL PROTEINS HYDROPHOBIC, SUGGESTING A POS HYDROPHOBIC, SUGGESTING A POS HYDROPHOBIC, SUGGESTING A POS HYDROPHOBIC, SUGGESTING A POS HYDROPHOBIC, SUGGESTING PHIS V LIPOPROTEIN ENVELOPE. THE ENV PROTEIN C AND MRNA.  SSNISS-PROT entry is copyright on the Swiss institute of B EUrophean Bioinformatics institute.	Dy non-prolit instructed and this statement les requires a license and an email to license bro; IPR002521; HCV_con Pro; IPR002521; HCV_con Pro; IPR002521; HCV_con Pro; IPR002521; HCV_con Pro; IPR01542; HCV_con Pro; IPR01542; HCV_con Pro; IPR01542; HCV_con Pro; IPR01560; HCV_mn; Inm; PD186062; HCV_mn; Inm; PD186062; HCV_mn; Inmembrane; Nonstructura MET 115 115
CVJ2 CCVJ2 27959; 27959; 11-AUG-19 6-OCT-20 enome po enome po fores) (G	Hepacivirus.  NGBI_TAXID=1111;  [1] SEQUENCE FROM N.A.  MEDLINE-9223023; Pubmed-1314459; Okamoto H., Kurai K., Okada S.I., Tanaka T., Fevuda S.I., Tanaka T., Fevuda S.I., Tanaka T., Fevuda S.I., Tanaka T., Fevuda F., M  Full-length sequence of a hepati homology to reported isolates: co genotypes."; Virology 188:331-341(1992).  -I-FUNCTION: THE SMALL PROTEINS HYDROPHOBIC, SUGGESTING A POS NS3 AND NS5 ANY PLAY A ROLE I -I-SUBNIT: THE VIRION OF THIS V LIPOPPOTEIN ROVELOPE. THE ENV PROTEIN AND GLYCOPROTEIN E. PROTEIN C AND MRNA.  This SWISS-PROT entry is copyrigh between the Swiss Institute of B the European Bioinformatics Institute	use by non-proint institution modified and this statement is entities requires a license agror send an enail to licensedish correction of send an enail to licensedish interpro; IPR002512; HCV_core. Interpro; IPR002519; HCV_core. Interpro; IPR0560; HCV_NS1; Interprom PP01560; HCV_NS1; Interpromensedistry interprospectation; Cortainsmembrane; Nonstructural printing INIT_MET Interpropression of CERAIN Interpropression of CERAIN Interpropression of CHAIN Interpression of CHAIN Interpression of CHAIN Interpression of CHAIN Interpress
RESULT POLG_H ID PPOLG_H AC P PO DT 0 O DT 1 DE G DE G O OC V	00000000000000000000000000000000000000	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC

us-09-491-146a-26.rsp

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JUL-1993 (Rel. 26, Last sequence update)
GLOT-2001 (Rel. 40, Last sequence update)
GLOCT-2001 (Rel. 40, Last annotation update)
GLOCT-2001 (Rel. 60, Last annotation update)
GLOCT-2001 (Rel. 61, GLOT-20, GLOT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-93019030; PubMed-1383400;
MEDLINE-93019030; PubMed-1383400;
MEDLINE-93019030; PubMed-1383400;
Abe K., Inchauspe G., Fujisawa K.;
"Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992);
-I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROFEIN MAND GLYCOPROTEIN E. THE BUNCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL)
                                                                      (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C BY CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 142; DB 1; Length 513;
Pred. No. 6.4e-13;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            943F31E3514CDEF3 CRC64;
                                                               N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            55704 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91.0%;
llarity 92.9%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D10688; BAA01530.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                             233
233
233
305
417
423
448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Local 26; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ¥,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVH4
ID POLG_HCVH4
AC Q01404;
                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
SEQUENCE
                                                                      CARBOHYD
                                                                                                                  CARBOHYD
                                                                                                                                                 CARBOHYD
                                       TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INIT_MET
ID DATE OF THE SECOND OF THE S
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@libs-sib.ch).
                                               MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66-OGT-2001 (Rel. 40, Last annotation update)
Envelope 91ycoprotein [Corpaid protein C (Core protein) (P22);
Envelope 91ycoprotein El (GP32) (GP35); Envelope 91ycoprotein E2
(GP68) (GP70) (NS1)] (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Abe K., Inchause G., Fullsawa K.;
"Genomic characterization and mutation rate of hepatitis C virus isolated from a patient who contracted hepatitis during an epidemic of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN EVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus (isolate HCV-KF) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein; Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
                          (POTENTIAL).
PROTEIN E (PO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                     AA135246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                     Score 142; DB 1; Lenral No. 6.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          520 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. 6.5e
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CAPSID
MATRIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A. MEDLINE-93019030; Pubmed-1383400;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002522; HCV_capsid
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Interpro; IPR002519; HCV_env.
Interpro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF015439; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                     ₹
                                                                                                                                                                                                                                                                                                                                                                                                                                                  91.0%;
ilarity 92.9%;
Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; D10687; BAA01529.1; -. JQ1925; JQ1925.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                  56499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                  520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
nes 26; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-31644;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLG_HCVHK
Q01403;
                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                    NON_TER
SEQUENCE
                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                              TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCVHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          THE PROPERTY OF THE PROPERTY O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
```

Fri Aug

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license & Sistement (See http://www.isb-sib.ch/announce/or send an email to license & Sistement (See http://www.isb-sib.ch/announce/or send an email to license & Sistement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                          ö
                           CAPSID PROTEIN ( POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
BY SIMILARITY.
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein); Matrix
E; Nonstructural
                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66nome polyprotein (Contains: Capaid protein C (Core protein); Matriprotein (Envelope protein M); Major envelope protein E; Nonstructura proteins (S1 and MS2) (Fragment).
Hepatitis C virus (isolate HC-15) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VITOLOGY 188:331-341(1992).

-1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MNAN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-9223023: PubMed-1314459;
Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
Okamoto H., Fuuda S., Tsuda F., Mishiro S.;
**Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                  Length 520;
   PROTEIN C
                                                                                                                                                                                                                                                                                                                 Score 142; DB 1; Length 52
Pred. No. 6.5e-13;
Pred. Totches 2; Indels
                                                                                                                                                                                                                                                                                                  MW; 1D2BD0A6FF27349B CRC64;
                  CELLULAR AMINOPEPTIDASE
   CAPSID
                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  737 AA.
                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
   FROM
                                                                                                                                                                                                                                                                                                                                                                                                           28
                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPORKPNRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_NSI.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D10075; BAA00969.1; -.
                                                                                                                                                                                                                                                                                                                                     91.08;
                                                                                                                                                                                                                                                                                                                                                      92.98;
                                                                                                                                                                                                                                                                                                  56476
                                                                                                                                                                                                                                                                                                                                                      Local Similarity 92.9
nes 26; Conservative
                                                                                                                                                                                                                                                                                                  520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genotypes."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVJ5
P27960;
                                                                                                                                                                                                                                                          CARBOHYD
NON_TER
SEQUENCE
                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                         CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                      TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                     Query Match
INIT_MET
                                                                                                                                       CARBOHYD
                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          8
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                      CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
POTENTIAL.
N-LINNED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       protein); Matrix E; Nonstructural
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VITOLOGY 188:331-341(1992).

-1- FUNCTION: THE SMALL PROFEINS NS2A, NS2B, NS4A AND NS4B ARE HYDOLOGY SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION. NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROFEINS: PROFEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDIINE-92230232; PubMed-1314459; Okamanoto K., Lizuka B., Okamoto H., Kurai K., Okada S.I., Yamanoto K., Lizuka B., Chanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                            POTENTIAL
                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL
                                                                                 ВУ
                                                                        protein.
REMOVED FROM CAPSID PROTEIN C
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                            Length 737
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                   CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core
protein (Bivelope protein M); Major envelope protein
proteins NS1 and NS2) (Fragment).
                                                                                            CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                91.0%; Score 142; DB 1; L
92.9%; Pred. No. 9.5e-13;
Nomatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                  81207 MW; 3AF699D82AD50131
                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     737 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus (fragment).
Viruses; ssRNA positive-strand viruses, no Hepatitus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28
                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
les 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVJ7
P27961;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genotypes.
                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                        FRANSMEM
                                                                              INIT_ME
                                                                                                        CHAIN
                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCVJ7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-505 (Rel. 23, Created)
01-505 (Rel. 23, Last sequence update)
01-506-1992 (Rel. 23, Last sequence update)
01-506-1992 (Rel. 23, Last sequence update)
01-506-1992 (Rel. 24, Last annotation update)
05-5003 (Rel. 42, Last annotation update)
05-5003 (Rel. 50-10014)
05-50
                                                                                                                                                                                                                                                                                                           CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                   protein.
REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91.0%; Score 142; DB 1; Length 737; 92.9%; Pred. No. 9.5e-13; Live 0; Mismatches 2; Indels
                                                                                                                                                                                                                        Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_core; 1.
                                                                                                                                                                                                                                                                     Nonstructural
                                                                                                   EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              I Similarity 92.99 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus.
NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                          192
334
347
196
196
209
209
4430
4430
448
                                                                                                                                                                                                                                                                     Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                 INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 6
POLG_HCVBK
                                                                                                                                                                                                                                                                                                                         CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license greement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     X MEDLINE-98227846; PubMed-956891;
X MEDLINE-98227846; PubMed-956891;
Yan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M., Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
T. "Complex of NS3 protease and NS4A peptide of BK strain hepatitis C T virus: a 2.2-A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-847(1998).

- I- FUNCTION: THE SWALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDOPHOBIC. SUGGESTING A POSSIBLE MEMBRANE-RLATED FUNCTION.
NS3 AND NS5 AMY PLAY A ROLLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 AMY PLAY A ROLLE IN THE VIRAL RNA REPLICATION.
C CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral presence of precursor polyprotein, commonly with Asp or Giu in the P6 precursor polyprotein.
C CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
MEDLINE-91140698; Pubmed-1847440;
Takamizawa A., Morl C., Fuke I., Manabe S., Murakami S., Fujita J.,
Onishi E., Andoh T., Yoshida I., Okayama H.;
"Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Love R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N., Moomaw E.W., Adachi T., Hostomska Z.; The crystal structure of hepatitis C virus NS3 proteinase reveals trypsin-like fold and a structural zinc binding site."; cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE, THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                     MEDLINE-96235224, PubMed-8647104,
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.
"Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by CAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE-97015088; Pubmed-8861916;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
HCV_NS2.
                                                                                                                                                  from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEAD.
                                                                                                                                                                                                                                                          SEQUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               26-JUN-00.
20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A38465; GNWVTC.
1A1Q; 25-MAR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1JXP; 14-JAN-98
1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10-APR-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PDB; 1CSJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BOHM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; InterPro; I
```

```
STRAND
SEQUENCE
STRAND
 a
                                                                                                                                                                                                                                                                                                                                                                                            CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                          Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                              REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                           CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                               InterPro; IPR002166; HCV_RGRP.
InterPro; IPR00166; HCV_RGRP.
InterPro; IPR001065; RNA_pol_DS_PS.
InterPro; IPR001094; RNA_pol_DS_PS.
InterPro; IPR001094; RNA_pol_PSVIr.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01559; HCV_NS2; 1.
Pfam; PF01560; HCV_NS3; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS5a; 1.
Pfam; PF01001; HCV_NS5a; 1.
Pfam; PF0100999; Viral_RGRP; 1.
Pfam; PF010999; Viral_RGRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2041
2077
2240
2529
2788
1035
1047
1063
                                                                                                                                                                                                                                                                                                     1081
1085
1087
1092
1094
1103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1076
                                                                                                                                                                                                                                                                  3D-structure,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                            INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAND
HELIX
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAND
                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HELIX
DORAN WAR AND CONTRACT OF THE CONTRACT OF THE
```

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein El (GP6) (RS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Japanese) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                    .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-91088550; PubMed-2175903;
Kato N., Hijkata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
"Molecular cloning of the human hepatitis C virus genome from
proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                Score 142; DB 1; Length 3010;
Pred. No. 4.5e-12;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                              327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                           91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
1112
1120
1120
1136
1136
11161
1161
1171
1174
11186
11186
11197
11197
                                                                                                                                                                                                                                                                                                                              3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11116;
                                                                                                        1149
11168
11168
11172
11175
11189
11189
11198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        [RNA](N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVJA
```

(POTENTIAL). (POTENTIAL). (POTENTIAL).

(POTENTIAL)

(POTENTIAL)

(POTENTIAL)

ö

ö

Length 3010;

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last aenocation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
GaP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.99..); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (Isolate H) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE=99154321; PubMed=9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonuclectide: the crystal structure provides insights into the mode
of unwinding";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -I- FUNCTION: PROTESSE N2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-IF FUNCTION: PROTESSE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.
-I- FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Prince A.M.; "Genomic structure of the human prototype strain H of hepatitis Virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-9733132; PubMed-9187654;
MEDLINE-9733132: Cable M., Hong Z., Kwong A.D., Le H.V.,
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                             ; DB 1; Lens
, 4.5e-12;
2; Indels
             N-LINKED (GLCNAC. . .) (POTENY 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches
                                                                                                                                                                                                                                                                                                                                    91.0%; Score 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPORKPNRNTNRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                         Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |||||||
PKPORKTKRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                        327017
                                                                                                                                                                                                                                                                                                                                                                     92.98;
                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Structure 6:89-100(1998)
                                                                            623
645
2041
2077
2240
2788
                                                                                                                                                                                                                                                                           Ä
                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus.
NCBI_TaxID-11108;
                                                                            623
645
2041
2077
2240
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (RNA) (N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLG_HCVH
P27958;
                                           CARBOHYD
CARBOHYD
                                                                                                           CARBOHYD
                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FTTTTTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
W Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Helicase; AFP-binding;
TINIT_MET 1 115 REMOVED FROM COPENTIAL;
CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).
CHAIN 192 383 MATOR ENVELOPE PROTEIN SI (POTENTIAL).
TCHAIN 194 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
TCHAIN 1007 1615 PROTESCHELICASE NSI (POTENTIAL).
TCHAIN 1007 1615 PROTESCHELICASE NSI (POTENTIAL).
TCHAIN 1661 2013 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
TCHAIN 1863 2013 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
TCHAIN 1863 2013 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
TCHAIN 1863 2014 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
TT CHAIN 1863 2014 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
TT RANSWEM 347 369 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
PROTEIN C AND MRNA.
SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002522; HCV_capsid.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002519; HCV_capsid.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR002519; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS4.
InterPro; IPR001490; HCV_NS4b.
InterPro; IPR001490; HCV_RS4b.
InterPro; IPR001656; HCV_RS4b.
InterPro; IPR001656; HCV_RS4b.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; HCV_capsid; IPR01543; HCV_capsid; IPR01543; HCV_capsid; IPR01543; HCV_capsid; IPR01540; HCV_NS1; IPR01006; HCV_NS3; IPR01006; HCV_NS4b; IPR01006; HCV_NS5a; IPR01006; HCV_NS4b; IPR01006; HCV_NS4b; IPR01006; HCV_NS4b; IPR01006; HCV_NS5a; IPR01006; HCV_NS4b; IPR01006; HCV_NS5a; IP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -LINKED
                                                                                                                                                                                                                                                                                                                                              EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00271; helicase_C; 1
Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2013
3010
369
11083
11107
11165
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
1213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                              HSSP; P26663; IJXP.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
```

LA TENERAL DE LA

Weber P.C.;

ပ

888888888888888888

us-09-491-146a-26.rsp

```
(POTENTIAL).
          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   327142 MW; 772CBB29CCD94753 CRC64;
                                                        (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                             (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                           (GLCNAC
         N-LINKED
                                                                                                                                                                                                                                                                   1224
12336
12336
12336
1247
1251
1277
1283
1302
1312
1313
1313
1313
1323
1333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1382
1389
1397
1410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1438
1450
1471
1481
1481
1481
1497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1343
1352
1360
1362
1368
1373
       CARBOHYD
CARBOHYD
CARBOHYD
                                                        HELLY
STRAND
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
HELLY
TURN
STRAND
TURN
STRAND
HELLY
                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation-the Buropean Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
-!- PTW: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
-!- SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39.
-!- SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001512; HCV_capsid.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR00166; HCV_NS5.
InterPro; IPR00266; HCV_RGP.
InterPro; IPR007099; RN_pol_DS_PS.
InterPro; IPR007099; RN_pol_DS_PS.
InterPro; IPR007094; NN_pol_DS_PS.
InterPro; IPR007094; NN_SS; I.
Pfam; PP01006; HCV_NS3; I.
Pfam; PP01001; HCV_NS4; I.
Pfam; PP00271; Hcl.case_C: I.
Pfam; PP00271; Hcv_NS4; I.
                                                                                                                                                                                                                                                                            EMBL; M67463; AAA45534.1; PIR; A36814; GNWVCH.
PDB; 1HEI; 25-NOV-98.
PDB; 1ANY, 16-FEB-99.
PDB; 1ANY, 17-JUN-98.
MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              191
383
746
809
1026
11657
1711
1972
2420
3611
363
1083
11165
11165
11165
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ProDom; PD186062; HCV_NSMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                        FRANSFAC; T04155; -.
InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

```
DORAN EFFERENCE ON STATE OF THE SECOND SEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     444B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-9204440; PubMed=1658196;
MEDLINE-9204440; PubMed=1658196;
MEDLINE-92044440; PubMed=1658196;
MEDLINE-92044440; PubMed=1658196;
Medida A., Miyakawa Y., Mayumi M.;
Mucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human cartier: comparison with reported isolates for conserved and divergent regions "."
J. Gen. Virol. 72:2697-2704(1991).
J. Gen. Virol. 72:2697-2704(1991).
HYDROPHOBIC, SUGGESTING A POESIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C. CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'
C. -- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                             01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 11, Last annotation update)
01-AUG-1992 (Rel. 11, Last annotation update)
03-FEB-2003 (Rel. 41, Last annotation update)
04-ORD (GROBA) (G
                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A SUBUNIT: THE VIRION OF THIS VIRUS IS A FWO PROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                     ö
                   Score 142; DB 1; Length 3011;
Pred. No. 4.5e-12;
0; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                          1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                              HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS3.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_env.
                91.0%;
llarity 92.9%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D00944; BAA00792.1; -. PIR; JQ1303; JQ1303. HSSP; P27958; JHEI. MERGOPS; S29.001; -. MERGOPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            interPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001490;
             Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=111113;
                                                                                                                                                                                                                                                                                                                        POLG_HCVJ6
ID POLG_HCVJ6
AC P26660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                               38844488
38844488
```

```
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; APT-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MW; F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                         CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 142; DB 1;
Pred. No. 4.6e-12;
0; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Created)
Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pro; IPR002868; HCV_NSSa.
rPro; IPR0001656; Hellcase_C.
erPro; IPR001059; RNA_POL_DS_PS.
erPro; IPR001095; RNA_POL_DS_PS.
erPro; IPR001095; RNA_POL_DS_PS.
erPro; IPR001094; RNA_POL_DS_PS.
erPro; IPR01394; HCV_core; I.
fam; PF011394; HCV_core; I.
fam; PF011394; HCV_core; I.
Pfam; PF011398; HCV_NS1; I.
Pfam; PF012904; HCV_NS2; I.
Pfam; PF012006; HCV_NS2; I.
Pfam; PF01001; HCV_NS4s; I.
Pfam; PF011001; HCV_NS4s; I.
Pfam; PF00101; HCV_NS4s; I.
Pfam; PF00101; HCV_NS4s; I.
Pfam; PF00101; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 92.9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                          115
191
191
191
1010
1619
1866
2017
3033
369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ¥,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 01-AUG-1992 (Rel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         649
1091
2038
2811
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVJ8
P26661;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 10
POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
```

9

```
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
         Genome polyprotein (Contains: Capsid protein ( (Core protein) (P22); Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP69) (GN31); Protein P7: Nonstructural protein NS2 (P21) (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS58 (P56); Nonstructural protein NS58 (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)]. Hepatitis C virus (Lapolate HC-J8) (HCV).
                                                                                                                                                                                                                                                                                          JIOOGY 188:331-341(1992).

-I- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-I- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

-I- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                  MEDLINE-92230232; PubMed-1314459; Okamoto K., Lizuka H., Tanaka Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka Bukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
28-FEB-2003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002522; HCV_capsid.

InterPro; IPR002519; HCV_core.

OR InterPro; IPR002519; HCV_core.

OR InterPro; IPR002519; HCV_env.

InterPro; IPR004109; HCV_NS1.

R InterPro; IPR004109; HCV_NS3.

InterPro; IPR00409; HCV_NS4a.

InterPro; IPR001499; HCV_NS5a.

InterPro; IPR002668; HCV_NS5a.

InterPro; IPR002668; HCV_NS5a.

InterPro; IPR007095; RNA_POL_DS_PS.

InterPro; IPR007095; RNA_POL_DS_PS.

InterPro; IPR007094; RNA_POL_DS_PS.

InterPro; IPR007094; RNA_POL_DS_PS.

InterPro; IPR007095; RNA_POL_DS_PS.

Pfam; PF01542; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Fram: PF01542; MCV_core; 1. Pfam: PF01542; MCV_core; 1. Pfam: PF01539; MCV_core; 1. Pfam: PF01539; MCV_NS1; 1. Pfam: PF02907; MCV_NS3; 1. Pfam: PF02907; MCV_NS3; 1. Pfam: PF01001; MCV_NS4s; 1. Pfam: PF01506; MCV_NS4s; 1. Pfam: PF01506; MCV_NS5s; 1. Pr0Dom: PF0098; VITAL_RGRP; 1. SMART; SMO0487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D10988; BAA01761.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A40250; GNWVJ8.
; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            THEI.
                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                            NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                     (RNA)(N).
                                                                                                                                                Hepacivirus
```

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
66-00me polyprotein (Contains: Capsid protein C (Core protein) (P22);
67-00 (GP68) (GP10) (NS1); Protein F1 (GP32) (GP32); Envelope glycoprotein E2 (GP68) (GP10) (NS1); Protein F2 Nonstructural protein NS2 (P21)
6C 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
6C 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructur
                                                                                           CELLULAR ANINOPERTIDASE.
CAPSID PROTEIN ( POTENTIAL).
MATRIX PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A ( POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Halicase; APP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                         RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY),
SIMILARITY),
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-91172826; PubMed-1848704;
Choo Q.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 142; DB 1; Length 3033;
Pred. No. 4.6e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                                                                                                        CHARGE RELAY SYSTEM (BY SCHARGE RELAY SYSTEM (BY CHARGE RELAY SYSTEM (BY S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNA)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity 92.9
nes 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                   733
1010
1619
1866
2017
3033
3033
1087
1111
1169
1241
1323
196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11104;
                                                                                                                                                                                                                                                                                                                                                   347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCV1
P26664;
                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                TRANSMEN
                                                                                                                                                                                                 CHAIN
CHAIN
CHAIN
                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
POLG_HCV1
```

```
88.5%;
                                                                                                                                                                                                                                                                                                                                                                                              88.38;
                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 89.3
nes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                        2041 207
2077 207
2240 224
2364 236
2789 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLG_HCVTW
P29846;
                                                          LSITE
BIND
                                                                                                CARBOHYD
                                                                                                                                                                CARBOHYD
                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                               Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 12
          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CELLULAR AMINOPEPTIDASE.

CAPSID PROTEIN (POTENTIAL).

MAJOR ENVELOPE PROTEIN E (POTENTIAL).

MAJOR ENVELOPE PROTEIN E (POTENTIAL).

NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).

NONSTRUCTURAL PROTEIN NSI (POTENTIAL).

NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).

NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).

NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
Gallegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,
Bradley D.W., Kuo G., Houghton M.;
Bradley D.W., Kuo G., Houghton M.;
Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1991).
- Frunction: THE SMALL PROFEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELAATED FUNCTION.
- SUGGESTING A POSSIBLE MEMBRANE RELAATED FUNCTION.
- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 postion, Cys or Thi in Pl and Ser or Ala in Pl'.
- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                         (RNA)(N).

SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                        SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002218; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001409; HCV_NS4a.
InterPro; IPR001409; HCV_NS4a.
InterPro; IPR001868; HCV_NS5a.
InterPro; IPR002868; HCV_NS5a.
InterPro; IPR00150; Helicase.
InterPro; IPR007095; RNA_Pol_DS_PS.
InterPro; IPR007095; RNA_Pol_DS_PS.
InterPro; IPR007094; RNA_Pol_DS_PS.
InterPro; IPR077094; RNA_Pol_PSvir.
Pfam; PF01543; HCV_Core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                             InterPro: IPR001410; DEAD.
InterPro: IPR002522; HCV_core.
InterPro: IPR002521; HCV_core.
InterPro: IPR002519; HCV_env.
InterPro: IPR002519; HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00271; helicase_C; 1
Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                    EMBL; M62321; AAA45676.1; -.
PIR; A39166; GNWVC3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF01006; HCV_NS4a;
PF01001; HCV_NS4b;
PF01506; HCV_NS5a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                             PDB; 1A1V; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          116
192
384
730
1007
1616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
```

```
Gaps
RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (KNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 138; DB 1; Length sv. Pred. No. 1.7e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MW; 65F8C9447FCE5AF9 CRC64;
                                                      (BY
(BY
(BY
                  POTENTIAL.
CHARGE RELAY SYSTEM (F
CHARGE RELAY SYSTEM (F
CHARGE RELAY SYSTEM (F
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKPNRNTNRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2240
2240
2364
2789
327197 M
```

Gaps

ö

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
MEDLINE=92295714; PubMed=1318627;
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Nakazawa T., Hijikata M., Ishimura Y., Shimotohno K.;
"Molecular cloning of hepatitis C virus genome from a single Japanese
carrier: sequence variation within the same individual and among
infected individuals.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VITUE RES. 23:39-53(1992).

-1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOSIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 postition, Cys or Thr in P1 and Ser or Ala in P1.

-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (RNA)(N).
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
    (POTENTIAL)
                                                                                                               (POTENTIAL)
                                                                                                                             (POTENTIAL)
                                                                                                                                         (POTENTIAL)
                    POTENTIAL)
                               POTENTIAL)
                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                             Length 3010;
                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                 MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                             Score 135; DB 1; Len
Pred, No. 4.6e-11;
                                                                                   (GLCNAC.
                                                                                                             (GLCNAC.
                                            (GLCNAC.
                                                                      (GLCNAC
                                                                                                                                          (GLCNAC
                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                   3010 AA
                                                                                                                                                                                                                                                                   2 KPORKPNRNTNRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                  6 KPQRKTKRNTNRRPQDVKFPGGGGIVG 32
               N-LINKED
86.5%;
92.6%;
                                                                                                                                                                                           Query Match
Best Local Similarity 92..
Best Local Similarity 92..
                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=31642;
                                                                                  623
645
2041
2077
2240
2529
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVJT
Q00269;
                                                                                CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                            CARBOHYD
                                                           CARBOHYD
                                                                        CARBOHYD
                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                      POLG_HCVJT
                                                                                                                                                                                                                                                                                                                                       RESULT 13
    g
                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA. SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pransmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BY SIMILARITY).
(BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MATRIX PROTEIN (POTENTIÁL).
MAJOR ENVELOPE PROTEIN E (P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELLULAR AMINOPEPTIDASE. CORE PROTEIN (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYS
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
                                                                                                                                                                                                                                                                                                        InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR00145; HCV_NS3.
InterPro; IPR000145; HCV_NS4a.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR002668; HCV_NS5a.
InterPro; IPR001266; HCV_RdRP.
InterPro; IPR007095; RNV_POL_DS_PS.
InterPro; IPR007095; RNA_POL_DS_PS.
                                                                                                                                                                   EMBL; M84754; -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                                                                                  HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF01538; HCV_NS2; 1. Pfam; PF01507; HCV_NS3; 1. Pfam; PF01006; HCV_NS4a; 1. Pfam; PF01001; HCV_NS4b; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF00271; helicase_C; 1. Pfam; PF00998; Viral_RGRP; 1. ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01543; HCV_caps1d; 1.
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; DEXDC;
                                                                                                                                                                                                                                                      InterPro; IPR001410;
                                                                                                                                                                                             1N64; 25-FEB-03,
1NS3; 08-APR-98
                                                                                                                                                                                                                                                                  IPR002522;
                                                                                                                                                                                A40244; GNWVTW
                                                                                                                                                                                                                        MEROPS; S29.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01539; PF01560; P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              347
1083
11107
11165
1230
1230
2233
2334
250
250
250
417
423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
INIT_MET
                                                                                                                                                                                                                                        . U39
                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                        MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               efam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
STRAIN=New Zealand; TISSUE-Brain, and Skeletal muscle;
MIDLINE-96161994; PubMed-8576224;
CSOTTOS C., Zolnierowicz S., Bako E., Durbin S.D., Depaoli-Roach A.A.;
"High complexity in the expression of the B' subunit of protein phosphatase 2A0. Evidence for the existence of at least seven novel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUBDINT: PPZA CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME, COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa CONSTANT REGULATORY SUBUNIT (PR65 or SUBUNIT A), THAT ASSOCIATES WITH A VARRETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS BY THE CAPE ASSOCIATE FAMILIES), THE 48 kDa VARTABLE REGULATORY SUBUNITS AND CELL SIGNALING MOLECULES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protein; Repeat.
8 X 2 AA APPROXIMATE TANDEM REPEATS OF Q-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     isoforms.";
J. Biol. Chem. 271:2578-2588(1996).

J. Biol. Chem. 271:2578-2588(1996).

-1- FUNCTION: THE REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
COMPARTMENT.
                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SH3 BINDING, CLASS I (POTENTIAL).
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
                                                                                                                                                                                                                                                                                                               Serine/threonine rot in phosphatase 2A, 56 kDa regulatory subunit, delta isoform (PP2A, B subunit, B' delta isoform) (PP2A, B subunit, B' delta isoform) (PP2A, B subunit, PRG1 delta isoform) (PP2A, B subunit, RS delta isoform) (PP2A, B subunit, RS delta isoform) (PP2A, B subunit, RS delta isoform) (PP2A, B subunit, B'-gamma).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CAUTION: NOMENCLATURE USED IN REF.1 REFERS TO PP2A B SUBUNIT B' GAMMA ISOFORM, WHICH IS CITED AS PP2A B SUBUNIT DELTA-PR61 ISOFORM IN LATER PUBLICATIONS.
                                                                                                                                                                                                                                                                                                                                                                                                                         Craniata; Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -i- SUBCELLUIAR LOCATION: Nuclear (By similarity).
-i- TISSUE SPECIFICITY: Highly expressed in brain.
-i- SIMILARITY: Belongs to the phosphatase 2A regulatory subunit
                                                                      ö
                                 Score 134; DB 1; Length 3010;
Pred. No. 6.4e-11;
0; Mismatches 3; Indels
326573 MW; 94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                         Bukaryota, Metazoa, Chordata, Craniata, Vertebrata, Eut.
Mammalia, Eutheria, Lagomorpha, Leporidae, Oryctolagús.
                                                                                                                                                                                                                          2A5D_RABIT STANDARD; PRT; 586 AA. 028653; 028655; 030-MAY-2000 (Rel. 39, Created) STANY-2000 (Rel. 39, Last sequence update) 16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                       1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                          Oryctolagus cuniculus (Rabbit).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; U38193; AAC48532.1; -.
                                 85.98;
                                                 Best_Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; U38195; AAC48534.1; -
InterPro; IPR002554; B56.
Pfam; PF01603; B56; 1.
Multigene family; Nuclear pr
3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=9986;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    507
SEQUENCE
                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                           RESULT 14.
2A5D_RABIT
                                                                                                                                         셤
                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2/E0 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-LIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
CHARGE RELAY SYSTEM ()
CHARGE RELAY SYSTEM ()
CHARGE RELAY SYSTEM ()
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                               nterPro; IPR007095; RNA_pol_DS_PS.
nterPro; IPR007094; RNA_pol_PSvir.
                                                                                                InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                           HCV_NS5a.
                                                                                                                                                                                                                                                                                                                          Pfam; PF01543; HCV_capaid; Pfam; PF01542; HCV_capaid; Pfam; PF01542; HCV_capaid; Pfam; PF01542; HCV_capaid; Pfam; PF01560; HCV_NS2; 1. Pfam; PF01006; HCV_NS3; 1. Pfam; PF01006; HCV_NS4a; 1. Pfam; PF01001; HCV_NS4a; 1. Pfam; PF01001; HCV_NS5a; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF00998; Viral_RARP; 1. Pr00998; Viral_RARP; 1. Pr000m; P0186062; HCV_NS1; 1.
                                                                                                                                                                                                       HCV_NS3.
HCV_NS4a.
HCV_NS4b.
       PIR; A45573; A45573.
PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            645
2041
2077
2240
2529
2788
                                                                                                                                                                              InterPro; IPR002518;
InterPro; IPR004109;
InterPro; IPR000745;
                                                                                                                                                                                                                                                             IPR002868;
                                                                                                                                                                                                                                           IPR001490;
                                                                               MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3D-structure
INIT_MET
                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                             interPro;
                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                             interPro;
                                                                                                                                                                                                                                                                                interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
```

qq

à

```
5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE—Colon, Eye, and Kidney.

WEDINE—22388257. PubMed—12477932.

MEDINE—22388257. PubMed—12477932.

Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

R. Altschul S.F., Zeeberg B., Buetow K.H., Schenfen C.M., Schuler G.D.,

Altschul S.F., Zeeberg B., Buetow K.H., Schenfer C.F., Bhat N.K.,

Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

Altschul S.F., Zeeberg B., Buetow K.H., Scheefer C.F., Bhat N.K.,

Altschul S.F., Zeeberg B., Buetow K.H., Scheefer T.E.,

Blackhenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

Rapleton M., Soares M.B., Bonaldo M.F., Carninci P., Prange C.,

Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

Raha S.S., Worley K.C., Hale S., Garcia A.M., Gabbs R.A.,

Richards S., Worley K.C., Hale S., Garcia A.M., Gabbs R.A.,

Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

Whiting M., Madan A., Young A.C., Schwutz J., Myers R.M.,

Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      subunits
PP2A to
                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                               014738; 000494; 000696; 015171; 01.00 MAY-2000 (Rel. 39, Created); 01.00 MAY-2000 (Rel. 39, Last sequence update); 01.00 MAY-2000 (Rel. 39, Last annotation update); 01.00 MAY-2000 (Rel. 42, Last annotation update); 01.00 MAY-2000 (Rel. 43, Last a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         or delta) of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. (ISOFORM DELTA-2), AND SEQUENCE OF 501-508; 550-559; 573-580 AND 584-601 (DELTA-1).

TISSUE-Bone marrow, and Brain cortex; MEDLINE-96159032; Pubmed-856619;

Tanabe O., Nagase T., Murakami T., Nozaki H., Usui H., Nishito Y., Tanabe O., Nagase T., Murakami T., Arakami T., Andecular cloning of a 74-kba regulatory subunit (B' or delta) of human protein phosphatase 2A.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens (Human). '
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             regulatory
                                                                                                               11;
                                                         Length 586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TISSUE-Fetal brain;
MEDIJNE-96355607; PubMed-8703017;
McCright B., Rivers A.M., Audlin S., Virshup D.M.;
The B56 family of protein phosphatase 2A (PP2A) regulatory encodes differentiation-induced phosphoproteins that target both nucleus and cytoplasm.";
                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE-Brain cortex;
MEDLINE-97324098; Pubhed-9180267;
Tanabe O., Gomez G.A., Nishito Y., Usul H., Takeda M.;
"Molecular heterogeneity of the cDNA encoding a 74 kDa subunit (8'' or delta) of human protein phosphatase 2A.
E149A309CDDA7495 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A. (ISOFORMS DELTA-1 AND DELTA-3).
                                                                                                                                                                                                    1:+| :| ::|:|| 30 PQPQSQP-PSSNKRPSNSTPPPTQLSKIKYSGGPQIV 65
                                                                                                                                                                     1 PKPQRKPNRNTNRRPQD------VKFPGGGQIV 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. (ISOFORMS DELTA-1 AND DELTA-2)
                                                                                                            3
                                                      DB 1;
                                                                                                                                                                                                                                                                                                                                                                                Ą
                                                   Score 55.5; D. Pred. No. 2.2; 8; Mismatches
                                                                                                                                                                                                                                                                                                                                                                             602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A. (ISOFORM DELTA-1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         h nucleus and cytoplasm.";
Biol. Chem. 271:22081-22089(1996).
                                                                                                                                                                                                                                                                                                                                                                          PRT;
68090 MW;
                                                   35.6%;
illarity 35.1%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEBS Lett. 379:107-111(1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEBS Lett. 408:52-56(1997).
586 AA;
                                                      Query Match
Best Local Similarity
Matches 13; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                HUMAN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PPP2R5D
                                                                                                                                                                                                                                                                                                                                        2A5D_HUMAN
ID 2A5D_H
AC Q14738
DT 30-MAY
DT 15-SEP
DE Serine
SO
```

RTH REPRESENT TO THE REPRESENT THE REPRESENT

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Genew, HGNC:9312; PPP2R5D.
GK; Q14738; -.
MIM; 601646; -.
GO:0005634; C:nucleus; TAS.
GO; GO:0006601; F:protein phosphatase type'2A, regulator acti. . ;; TAS.
GO; GO:0007399; P:neurogenesis; TAS.
InterPro; IPR002554; B56.
                                                                                                                                                                                                                                                         SUBBUILT: PPAR CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME, COMPOSED OF A 36 kDa CATALYTIC SUBBUIT (SUBBUIT C). AND A 65 kDa CONSTANT REGULATORY SUBBUIT (PR65 OR SUBBUIT A), THAT ASSOCIATES WITH A VARIETY OF REGULATORY SUBBUITS. PROTEINS THAT ASSOCIATE WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBBUITS BY CHREE STAFFILES), THE 48 kDA VARIABLE REGULATORY SUBBUIT; AND CELL SIGNALING MOLECULES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SH3 BINDING, CLASS I (POTENTIAL).
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
Missing (in isoform Delta-3).
/Frida-VSP_005110.
Missing (in isoform Delta-2).
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schorch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

-!- FUNCTION: THE B REGULATORY SUBUNT MIGHT MODULATE SUBSTRATE SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ISOId-014738-3; Sequence=VSP_005110;
--- TISSUE SPECIFICITY: ISOFORM DELTA-2 IS WIDELY EXPRESSED. ISOFORM DELTA-1 IS HIGHLY EXPRESSED IN BRAIN.
--- INDUCTION: BY RETINOIC ACID; IN NEUROBLASTOMA CELL LINES.
--- PTM: AT LEAST ISOFORM DELTA-1 IS PHOSPHORYLATED ON SERINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8 x 2 AA APPROXIMATE TANDEM REPEATS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -1- SIMILARITY: Belongs to the phosphatase 2A regulatory subunit B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR IN INTERPHASE, NUCLEAR DURING MITOSIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Phosphorylation; Alternative splicing; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; Length 602;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /FTId=VSP_005111.
F15F71AF4E565387 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Event-Alternative splicing; Named isoforms-3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 53.5; DI
Pred. No. 4.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IsoId=014738-2; Sequence=VSP_005111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IsoId=Q14738-1; Sequence-Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AB000634; BAA20381.1; -. AB000635; BAA20382.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           D78360; BAA11372.1; -. BC001095; AAH01095.1; -. BC001175; AAH01175.1; -. BC010692; AAH10692.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; L76702; AAB69751.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       34.3%;
31.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                69991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ALTERNATIVE PRODUCTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  530
565
116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR; S68686; S68686
Genew; HGNC:9312; Pi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF01603; B56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                602 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Multigene family.
DOMAIN 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Name-Delta-2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Name-Delta-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Name-Delta-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   82
                                                                                                                                                                                                                                          COMPARTMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESIDUES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
```

us-09-491-146a-26.rsp

Matches 13; Conservative 8; Mismatches 6; Indels 15; Gaps δ

;

1 PKPQRKPNR-----VKTPGGGQIV 27 | :| | :| | | | | 40 PQPQPQAQSQPPSSNKRPSNSTPPPTQLSKIKYSGGPQIV 81

qq

Search completed: August 7, 2003, 11:19:59 Job time: 5.90909 secs

Q8v7wl hepatitis

**08V7W1** 

```
12
 191
 94.9
 148
17
                                                                                                     August 7, 2003, 11:05:41; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                           830525
              GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                         tal number of hits satisfying chosen parameters:
                                                                                                                                                                    US-09-491-146A-26
156
1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                           830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                           OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_plant:*
sp_rodent:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_archea:*
sp_bacteria:*
sp_fung1:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                        Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                  SPTREMBL_23:*
                                                                                                                                                                    Title:
Perfect score:
                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Database :
                                                                                                                                                                                                     Sequence:
                                                                                                                                                                                                                                                                              Searched:
                                                                                                          Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		. u	O	U	υ	ပ	υ	Ų	υ	υ	O	O	U	U	U	U	O
	ion	hepatitis	hepatitis	hepatitis	hepatitis	_	_	_	hepatitis	_	_	hepatitis	hepatitis	hepatitis	hepatitis	_	_
	Description	08v7s2	08v7t0	Q8v7s4	081291	081p76	Q8v7x8	08v7v4	Q8v7x5	Q8v7q3	08v7k4	Q8v7m4	Q8v711	Q8v7y3	08v7n6	Q8v7y0	08v7k9
	ID	084752	Q8V7T0	Q8V7S4	081291	Q8JP76	Q8V7X8	Q8V7V4	Q8V7X5	08V703	Q8V7K4	Q8V7M4	Q8V7L1	Q8V7Y3	<b>08V7N6</b>	Q8V7Y0	Q8V7K9
	DB	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
	Query Match Length DB	191	191	191	119	191	191	191	191	191	191	191	191	191	191	191	191
æ	Query	96.2	96.2	96.3	94.9	94.9	94.9	94.9	94.9	94.9	94.9	94.9	94.9	94.9	94.9	94.9	94.9
	Score	150	150	150	148	148	148	148	148	148	148	148	148	148	148	148	148
	Result No.	-	7	e	4	S	9	7	80	თ	10	11	12	13	14	15	16

|--|

Gaps

ö

119 AA.

Length 191;

음

```
Hotta H., Handajani R., Ingelusida M., Soemarto W., Doi H.,
Miyajima H., Homma M.;
Subtype analysis of hepatitis C virus in Indonesia on the basis of
NS5b region sequences.;
J. Clin. Microbiol. 32:3049-3051(1994).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GIXCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MANA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyak S.J., Miller K., McArdle S., Gale M.J. Jr., Tenoever B., Hiscott J., Gretch D.R.; "Quasispecies Dependent Activation of the Innate Interferon System By
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94.9%; Score 148; DB 12; Length 119; 96.4%; Pred. No. 8.2e-14; 1.ve 0; Mismatches 1; Indels
                                                                                  640D7CA82765EF34 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               119 AA; 13510 MW; F6836FBFA5E5EB2A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                   Score 150; DB 12;
Pred. No. 7e-14;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PKPQRKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                 5 PKPQRKPKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPORKPNRNTNRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                     1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-0CT-2002 (TrEMBLrel. 22, Created)
01-0CT-2002 (TrEMBLrel. 22, Last sequ
01-MAR-2003 (TrEMBLrel. 23, Last seno
                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Core protein (Genome polyprotein). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-Td-34/92;
MEDLINE-95189942; PubMed-7883898;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro, IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                  20744 MW;
                                                                                                                                   Query Match
Best Local Similarity 96.2%;
Matches 27; Conservative
     Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity 96.4 ies 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                               191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11103;
                                                      191
                        Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-coreD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8JP76;
                                                                                                                                                                                                                                                                                                                                                                                                                       081291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08JP76
                                                                                                                                                                                                                                                                                                                                                                   RESULT 4
Q81291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08JP76
                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AB062194; BAB83388.1; -InterPro; IPR002522, HCV_capsid.

InterPro; IPR002521; HCV_capsid.

Pfam; PF01543; HCV_capsid, 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 150; DB 12; Length 191; Pred. No. 7e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     191 191
191 AA; 20740 MW; 64151CA83695EF34 CRC64;
                                                                                                                                                                                            or man-zouz (TrEMBLrel. 20, Last sequence update) 01-0CT-2002 (TrEMBLrel. 22, Last annotation update) core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ol-MAR-2002 (TrEMBLrel. 20, Last sequence update) 01-0CT-2002 (TrEMBLrel. 22, Last annotation update) Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 96.2%; Score 150; DB Best Local Similarity 96.4%; Pred. No. 7e-1 Matches 27; Conservative 0; Mismatches
PKPQRKPKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL, AB062200; BAB83394.1; -.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-MAR-2002 (TrEMBLrel. 20, 01-MAR-2002 (TrEMBLrel. 20, 01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                            01-MAR-2002 (TrEMBLrel. 20,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE-Liver;
Kato N.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  lesions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              lesions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q8V7S4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kato N.
                                                                                                                             Q8V7T0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q8V7S4
                                                                         RESULT 2
(080770
(080770
AC (08077
AC (08077
AC (08070
DT (01-M
DT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 3
```

08V7S4 ID Q

셤

ö

Gaps

ö

191 AA.

```
Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROPEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROFIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AB062025; BAB83364.1; -.

InterPro; IPR002521; HCV_capsid.

Pfam; PF01543; HCV_capsid.

Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (MAY-2001) to the EMBL/GenBank/DD3J databases.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROPEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROPEINS: PROTEIN MAND GLYCOPROPEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AB062004; BAB83343.1; - InterPro; IPR002521; HCV_capsid.

InterPro; IPR002521; HCV_capsid.

InterPro; IPR002521; HCV_core.

Pfam; PF01543; HCV_core; 1.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  'Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                           Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER 191 191 SEQUENCE 191 AA; 20694 MW; 2F4BD80E7BFE8E9A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191 191 191 191 191 AA; 20728 MW; 23C8ABC1783BB49A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAR-2002 (TrEMBLrel. 20, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94.9%; Score 148; DB 12;
ilarity 96.4%; Pred. No. 1.4e-13;
Conservative 0; Mismatches 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 148; DB 12;
Pred. No. 1.4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKPNRNTNRRPQDVKFPGGCQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKPQRKTNRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    94.98;
96.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (TrEMBLrel. 20, (TrEMBLrel. 20,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
es 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11103;
                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  rissue-Liver;
                                                                                                                                                         TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAR-2002
01-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                         lesions.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     lesions.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kato N.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8V7X5;
                                                                                                                                                                                 Kato N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q8V7X5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
           STATE TO COURT TANK SERVICE SOLUTION OF THE SERVICE SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Lesions... | Les
  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94.9%; Score 148; DB 12; Length 191; 96.4%; Pred. No. 1.4e-13; 1ve 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                         Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein.
NON_TER 191 191
SEQUENCE 191 AA; 20718 MW; 23C8A9ACE83B849A CRC64;
                                                                                                                                                                                                                                                                                                                                                   191 AA; 20797 MW; 3632F808598A348A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-WAR-2002 (TrEMBLrel. 20, Created)
01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                      Score 148; DB 12;
Pred. No. 1.4e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PKPORKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                            94.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IISSUE-Liver;
                                                                                                                                                                                                                                                                                                                               Polyprotein.
SEQUENCE 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Kato N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q8V7X8
Q8V7X8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      087774;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q8V7V4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08V7V4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
              à
```

Gaps

ö

ö

Gaps

ö

4

```
lesions.";
Submitted
      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                 Kato N.;
                                                                                                                           08V7M4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q8V7L1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q8V7L1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q8V7L1
                                                             ò
                                                                               a
                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                             "Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, ABM62270; BAB83464.1;

InterPro; IPR002522; HCV_capsid.

InterPro; IPR002521; HCV_capsid.

Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                   Score 148; DB 12; Length 191;
Pred. No. 1.4e-13;
0; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                        191 191
191 AA; 20794 MW; 36DD3CBAE63B8532 CRC64;
                                     01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-0CT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAR-2002 (TrEMBLrel. 20, Created)
1-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-CCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                     191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                         191 AA.
                                                                                                                                                                                                                                                                                                                                                       1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                  S PKPQRKTNRNTNRRPQDVKFPGGGOIVG 32
                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                   94.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAR-2002 (TrEMBLrel, 20,
                                                                                                                                                                                                                                                                                                         Ouery Match
Best Local Similarity 90.4.
The 27; Conservative
                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                    Hepatitis C virus
                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
TISSUE-Liver;
                                                                                                                NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                             Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                            TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
NON_TER
                                                                                                                                                                        lesions."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Kato N.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 lesions.
                                                                                                                                                                                                                                                                                                                                                                                                                        08V7K4
RESULT 9
Q8V7Q3
                      SE TA BOR SE S
                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                        14451
14451
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                               Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
-- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROFEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                        Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               94.9%; Score 148; DB 12; Length 191; 96.4%; Pred. No. 1.4e-13; Live 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             lesions.";
Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
191 AA; 20718 MW; 23C8A9ACE83B849A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191
20746 MW; 23DB0B26D33B849A CRC64;
                                                                                                                                                                                                                                                                                 01-WAR-2002 (TrEMBLrel. 20, Created)
01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                  94.9%; Score 148; DB 12;
96.4%; Pred. No. 1.4e-13;
11ve 0; Mismatches 1;
                                                                                                                                                                                                                                                  191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   191 AA.
                                                                                                            1 PKPORKPNRNTNRRPODVKFPGGGQIVG 28
                                                                                                                                1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTNRNINRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AB062250; BAB83444.1; -.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 96.4%;
nes 27; Conservative
                                Query Match
Best Local Similarity 96.44
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        191 1
191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE-Liver;
```

us-09-491-146a-26.rspt

```
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                            .
G
                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                         Gaps
   COVERED BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

-!- SUBMNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; A8061996; BAB83335.1; -
Interpro; IPR002522; HCV_capsid.
-I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL, AB062263; BAB83457.1; -

InterPro; IPR002521; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
From PF01543; HCV_core.
Pfam; PF01542; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           'Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                              94.9%; Score 148; DB 12; Length 191; 96.4%; Pred. No. 1.4e-13; Live 0; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191
20718 MW; 23C8A9ACE83B849A CRC64;
                                                                                                                                       SEQUENCE 191 AA; 20730 MW; 23C8A9ACE82A749A CRC64;
                                                                                                                                                                                                                                                                                                                     01-MAR-2002 (TTEMBLrel. 20, Created)
01-MAR-2002 (TTEMBLrel. 20, Last sequence update)
01-OCT-2002 (TTEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-WAR-2002 (TrEMBLrel. 20, Created)
01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94.9%; Score 148; DB 12;
96.4%; Pred. No. 1.4e-13;
tive 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Core protein (Genome polyprotein) (Fragment) CORE.
                                                                                                                                                                                                                                                                                                    191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                               1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                             PKPORKINRNINRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 96.41
Matches 27; Conservative
                                                                                                                                                                           Best Local Similarity 96.4 Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=11103;
                                                                                                                          191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE-Liver;
                                                                                                              Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           lesions."
                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08V7N6
08V7N6;
                                                                                                                                                                                                                                                                                                    Q8V7Y3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Kato N.
                                                                                                                                                                                                                                                                          RESULT 13
Q8V7Y3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14
                                                                                                                                                                                                                                                                                                     ö
  ò
                                                                                                                                                                                                                                     5
```

```
"Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                            lesions.";
submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
-!- SUBMITT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AB062238; BAB83432.1; -.

InterPro; IPR002522; HCV_caps1d.

InterPro; IPR002522; HCV_caps1d.

InterPro; IPR002521; HCV_core.

Pfam; PF01543; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Submitted (MAX-2001) to the EMBL/GenBank/DDBJ databases.

-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AB061999; BAB83338.1; -
InterPro; PR002522; HCV_capsid.

InterPro; IPR002521; HCV_capsid.

InterPro; IPR002521; HCV_core.

Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             'Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match 94.9%; Score 148; DB 12; Length 191; Best Local Similarity 96.4%; Pred. No. 1.4e-13; Matches 27; Conservative 0; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                           191
20730 MW; 23C8A9ACE82A749A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        191 191 WW; 23C8A9ACE83B849A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAR-2002 (TrEMBLrel. 20, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 148; DB 12;
Pred. No. 1.4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
94.9%; Score 148; DE
Best Local Similarity 96.4%; Pred. No. 1.46
Matches 27; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last seq
01-0CT-2002 (TrEMBLrel. 22, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                 Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          lesions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kato N.;
                                                      Kato N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8V7Y0
```

Search completed: August 7, 2003, 11:18:59 Job time : 26.6364 secs

us-09-491-146a-26.rspt

```
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 1, Appliance 23, Appliance 26, Appliance 27, Appliance 26, Appliance 26, Appliance 27, Appliance 10, Appliance 198, Appliance 198, Appliance 198, Appliance 123, Appliance 134, A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 26, Appl
                                                                                                                                                                                                                                                 August 7, 2003, 11:07:41; Search time 10.5455 Seconds (without alignments) 112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 6, Ap
Sequence 36, A
Sequence 2, Ap
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-836-075A-10
US-08-635-886C-198
US-08-635-886C-232
US-08-836-075A-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-380-160-2
US-09-389-756-1
US-08-380-160-1
US-07-946-054-9
US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-501-195-2
US-08-836-075A-46
US-08-836-075A-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PCT-US93-08638-9
PCT-US94-07088-23
PCT-US95-13660-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       156
1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-530-550-3
US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PCT-US95-13552-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-324-977-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                         - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gapop 10.0 , Gapext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Minimum DB seq length: 0 Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-491-146A-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             %
Ouery
Match I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Perfect score:
                                                                                                                                                                         OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Searched:
                                                                                                                                                                                                                                                              Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Result
```

9991 9911 9999 9999 9999 9999 9999 999	THE SA	;0
«C.		ss C
	MAK	Gaps
sednence sed	O9 . RO	, ,
sedneuce sed	METHOD	Length 28; Indels
28 142 91.0 190 1 US-07-681-701-16 29 142 91.0 190 3 US-08-078-271B-1 30 142 91.0 191 2 US-08-090-665A-155 31 142 91.0 191 2 US-08-290-665A-155 33 142 91.0 191 2 US-08-290-665A-158 34 142 91.0 191 2 US-08-290-665A-158 35 144 91.0 191 2 US-08-290-665A-161 36 142 91.0 191 2 US-08-290-665A-161 37 142 91.0 191 2 US-08-290-665A-161 38 142 91.0 191 2 US-08-290-665A-164 39 142 91.0 191 2 US-08-290-665A-165 40 142 91.0 191 2 US-08-290-665A-165 41 142 91.0 191 2 US-08-290-665A-165 42 142 91.0 191 2 US-08-290-665A-165 43 142 91.0 191 2 US-08-290-665A-168 44 142 91.0 191 2 US-08-290-665A-168 45 142 91.0 191 2 US-08-290-665A-168 46 142 91.0 191 2 US-08-290-665A-168 47 142 91.0 191 2 US-08-290-665A-170 48 142 91.0 191 2 US-08-290-665A-170	RESULT 1 19-087-26 19-0897-26 19-0897-26 19-0897-26 19-0897-26 19-0897-26 19-087-26 19-087-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-27 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28 19-08-28	Query Match 100.0%; Score 156; DB 3; Leng Best Local Similarity 100.0%; Pred. No. 1e-14; Matches 28; Conservative 0; Mismatches 0; Inc
	S S S S S S S S S S S S S S S S S S S	Omz

```
SYNTHETIC POLYPEPTIDES BELONGING TO THE HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY FOR DETECTING THE LATTER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 43;
                                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,846
FILING DATE: 09-FEB-1998
CLASSIFICATION NUMBER: US/09/020,846
FILING APPLICATION DATA:
APPLICATION NUMBER: 1999
FILING APPLICATION NUMBER: 10-FEB-1997
PRIOR APPLICATION NUMBER: 10-FEB-1997
APPLICATION NUMBER: 19-FEB-1996
ATTON NUMBER: 09-FEB-1996
ATTON NUMBER: 25,258
REGISTRATION NUMBER: 25,258
REGISTRATION NUMBER: 25,258
REGISTRATION NUMBER: 053466/0225
TELEPHONE: (202) 672-5300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 142; DB 4; Ler
Pred. No. 1.3e-12;
    APPLICANT: CHIBA, Yukie
APPLICANT: YAGI, Shintaro
APPLICANT: HASEGAWA, AKIRA
TITLE OF INVENTION: CHIMERA ANTIGEN PEPTIDE
NUMBER OF SEQUENCES: 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                          COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYREDE Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 2, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
APPLICANT: DALBON, Pascal
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POLY
TITLE OF INVENTION: HEPATITIS C VI
TITLE OF INVENTION: HEPATITIS C VI
TITLE OF INVENTION: FOR DEFECTING
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STREET: P.O. BOX 19928
CITY: Alexandria
STATE: VA
                                                                                                     CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W
CITY: Washington
STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          43 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                 APPLICANT: DALBON, Pascal
APPLICANT: DALBON, Pascal
APPLICANT: DALBON, Pascal
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHEIR
TITLE OF INVENTION: FOR DETECTING THE LATTER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
STRRET: P.O. Box 19928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91.0%; Score 142; DB 3; Length 34; 92.9%; Pred. No. 9.9e-13; tive 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION: 530
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AEBNT INFORMATION:
NAME: BETLIGG, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
TELEPHONE: (703)936-6400.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: Human Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Sequence 36, Application US/09020846; Patent No. 6322965; GENERAL INFORMATION:
                                                                                                                                          Sequence 6, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YAMAGUCHI, Kenjiro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FRAGMENT TYPE: N-terminal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INFORMATION FOR SEQ ID NO:
SEQUENCE CHRRACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acids
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match
Best Local Similarity 92.9°
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZIP: 22320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
US-09-020-846-36
                                                                                                                         US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-380-160-6
ô
                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ω
```

Gaps ö

```
Sequence 1, Application US/08380160

Patent No. 6235284
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: DALBON, PASCAL
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: FOR DETECTING THE LATTER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
    AUTHORS: Ogata, N. et al. TITLE: Nucleotide Sequence and Mutation Rate of the H Strain Patent No. 6576240
                                                                                                                                                                                                                                            Query Match 91.0%; Score 142; DB 4; Length 44; Best Local Similarity 92.9%; Pred. No. 1.3e-12; Matches 26; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                       1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: BETINGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
                                                                TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 88
PAGES: 3392-3396
DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OKGANISM: Human Hepatitis C Virus STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 45 amino acids
                                                                                                                                                                                ; RELEVANT RESIDUES: 2 TO 45 US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: peptide HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 6
                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: JOLIVET, MICHEL
APPLICANT: DALGNU, FRANCOIS
APPLICANT: DALGNU, PASCAL
APPLICANT: DALGNU, PASCAL
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILE REFERENCE: 103959
CURRENT FILING DATE: 1999-09-07
EARLIER APPLICATION NUMBER: PCT/FR98/00442
EARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 142; DB 3; L
Pred. No. 1.3e-12;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4 PKPQKKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ANTI-SENSE: NO
FRACHENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Human Hepatitis C Virus
STRAIN: H77
                                                                                                                                      CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAX-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sequence 1, Application US/09389756 Patent No. 6576240
                                                                                                                                                                                                                                                   NAME: Berridge, William P. REGISTRATION NUMBER: 30,024 REFERENCE/DOCKET WUMBER: WPIPELECOMMUNICATION INFORMATION: (703)836-6400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Hepatitis C virus PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                        INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: amino acids
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                              TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; OTHER INFORMATION: /
; OTHER INFORMATION: p.
; OTHER INFORMATION: th
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY: Peptide LOCATION: 1..44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ENERAL INFORMATION:
                                                                                                                        FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        :ULT 5
-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
No. 5639594el Linear and Branched Peptides Effective
in Diagnosing and Detecting No. 5639594-A, No. 5639594-J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 3, Application US/08530550
Patent No. 5736321
GENERAL INFORMATION:
APPLICANT: Hosein, Barbara
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Peptides Effective for Diagnosis and
TITLE OF INVENTION: Detection of Hepatitis c Infection
                                                                                                                                                                                                                              MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN RELEASE #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,947
FILING DATE: 19930628
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 946,054
FILING DATE: 15-SEP-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 142; DB 1;
Pred. No. 1.8e-12;
0; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                TITLE OF INVENTION: No. 56395:
TITLE OF INVENTION: in Diagno:
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: W11SON, M. LISA
RECISTRATION NUMBER: 34045
REFERENCE/DOCKET NUMBER: 200
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (516)273-2828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEFAX: (5/5/3/3/3/2021)
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                         ZIP: 11788
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               amino acid
                                                                                                                                    Hauppauge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET: 25 LUC
                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-083-947-23
                                                                                                                                                                           COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Wang, Chang Xi
APPLICANT: Wang, Chang Xi
APPLICANT: Hosein, Barbara H
TITLE OF INVENTION: No. 5582968-1 Branched Hybrid and Cluster
TITLE OF INVENTION: Peptides Effective in Diagnosing and Detecting No. 5582968-A,
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: United Biomedical Inc.
STREET: 25 Davids Dr.
                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ô
                      /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                   ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                  Ouery Match
91.0%; Score 142; DB 3; Length 45;
Best Local Similarity 92.9%; Pred. No. 1.3e-12;
Matches 26; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 142; DB 1; Length 61;
Pred. No. 1.8e-12;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/946,054
FILING DATE: 15-SEP-1992
CLASSIFICATION: 435
                                                                                                                                                                                               1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                    1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 PKPQRKTKRNINRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZIP: 11788
COMPUTER REARABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 23, Application US/08083947
Patent No. 5639594
GENERAL INFORMATION:
APPLICANT: Wang, Chang Y1
APPLICANT: Hosein, Barbara
                                                                                                                                                                                                                                                                                                                                    Sequence 9, Application US/07946054
Patent No. 5582968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME: WILSON, M. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/DOCKET NUMBER: 2000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-273-2828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 92.9s;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEFAX: 516-273-1717
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS
COCATION: 1..45
COTHER INFORMATION:
COTHER INFORMATION:
COTHER INFORMATION:
US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOLECULE TYPE: peptide US-07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: Hauppauge
STATE: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 USA
                                                                                                                                                                                                                                                                                                           US-07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 8
                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

Gaps

```
Sequence 9, Application PC/TUS9308638
Sequence 9, Application PC/TUS9308638
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
NO.B Hepatitis
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
ADDRESSEE: UNITED BIOMEDICAL INC.
                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 142; DB 5; Length 61;
Pred. No. 1.8e-12;
0; Mismatches 2; Indels
                                                         Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Wang, Chang Yi
APPLICANT: Hoseln, Barbara
TITLE OF INVENTION: Novel Linear And Branched
TITLE OF INVENTION: Peptides Effective In
TITLE OF INVENTION: Diagnosing And Detecting
TITLE OF INVENTION: Non-A, Non-B Hepatitis
NUMBER OF SEQUENCES: 23
                                                         Score 142; DB 1;
Pred. No. 1.8e-12;
0; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                            1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 23, Application PC/TUS9407088 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                 E: UNITED BIOMEDICAL INC. 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE: Floppy disk COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLESSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: M. Lisa Wilson
REGISTRATION NUMBER: 34,045
REFERENCE/DOCKET NUMBER: 9055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516_273-2828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                           91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
                                                           Query Match
Best Local Similarity 92.99
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ZIP: 11788
COMPUTER REDABLE FORM:
MEDIUM TYPE: Flore:
COMPUTED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08638-9
    Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hauppauge
New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PCT-US94-07088-23
; TOPOLOGY:
US-08-262-037-26
                                                                                                                                                                                                                                                                   PCT-US93-08638-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 12
                                                                                                                                                                                                                                                RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                               δ
                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Chang Y1 Wang and Barbara Hosein
TITLE OF INVENTION: SYNTHETIC PEPTIDES SPECIFIC FOR
TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
TITLE OF INVENTION: INFECTION AND PREVENTION THEREOF AS VACCINES
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSE: MORGAN & FINNEGAN
STREET: 345 PARK AVE.
CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                   ;
0
                                                                                                                                                                                                                                                                                             Length 61
                                                                                                                                                                                                                                                                        Score 142; DB 1; Le
Pred. No. 1.8e-12;
----hes 2;
                                                                                                                                                                                                                                                                                                                                     0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 29,323
R: 1151-4043 US3
                                                                                                                                                                                                                                                                                                                                                                         1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                    4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/262,037
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: 07/558,799
FILING DATE: 06-501y-1990
FILING DATE: 16-501y-1990
APPLICATION NUMBER: 07/510,153
FILING DATE: 16-API1-1990
ATTORNEY/FORDY INFORMATION:
NAME: Maria C. H. Lin
REGISTRATION NUMBER: 29,323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/719,819
FILING DATE: 24-June-1991
APPLICATION NUMBER: 07/667,275
FILING DATE: 11-Mar-1991
APPLICATION NUMBER: 07/651,735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 26, Application US/08262037
Patent No. 5747239
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE/DOCKET NUMBER: 11
TELECOMMUNICATION:
TELEPHONE: 212-758-4800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
                           REGISTRATION NUMBER: 34,045
REPERENCE/DOCKET UNBER: 20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516)273-2828
                                                                                                                                                                                                                                                                                           Query Match 91.0%;
Best Local Similarity 92.9%;
Matches 26; Conservative C
                                                                                  TELEPHONE: (516)273-2828
TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIUM TYPE: FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH: 61 amino acids
                                                                                                                                                                         LENGTH: 61 amino acids
                                                                                                                                                                                                                             , MOLECULE TYPE: peptide US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Amino acid
                                                                                                                                                                                         TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: Amino
STRANDEDNESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE: N
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                 à
```

Gaps

```
GENERAL INFORMATION:

APPLICANT: MARRIENS, GERT
APPLICANT: MARRIENS, LIEVER
APPLICANT: STUVYER, LIEVER
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOST:
TITLE OF INVENTION: AGENTS
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                       ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                            Score 142; DB 5; Length 61;
Pred. No. 1.8e-12;
                                                                                                                                                                                                                                                                                                                                                                    2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ZIP: 77210-4433
ZIP: 77210-4433
ZUP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-POS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,075A
FILING DATE: 23 Oct 1995
FILING DATE: 23 Oct 1995
FILING DATE: 21 Oct 1995
FILING DATE: 21 Oct 1994
PRIOR APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
PRIOR APPLICATION NUMBER: EP 9487016.9
FILING DATE: 28 Jun 1995
ATTORIEZ/AGENT IRPORMATION:
AMARIEZ/AGENT IRPORMATION:
AMARIEZ/AGENT IRPORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 142; DB 3; I
Pred. No. 2.2e-12;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                     Query Match 91.0%; Score 142; DE Best Local Similarity 92.9%; Pred. No. 1.86 Matches 26; Conservative 0; Mismatches
           REGISTRATION NUMBER: 29,323
REGISTRATION NUMBER: 1151-4118PC
TELECOMUNICATION INFORMATION:
TELEPRAX: (212)751-689
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acids
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                    4 PKPQKKTKKNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INNS: 004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Sequence 10, Application US/08836075A
; Patent No. 6180768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775 REFERENCE/DOCKET NUMBER: IN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 91.0
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P.O. BOX 4433
                                                                                                                                                                                                                                              ; MOLECULE TYPE: peptide
PCT-US95-13660-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 14
US-08-836-075A-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-836-075A-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 3, Application PC/TUS9513660
GENERAL INFORMATION:
APPLICANT: Hosein, Barbara
APPLICANT: Wang, Chang Yi
TTLE OF INVENTION: Peptides Effective for
TITLE OF INVENTION: Diagnosis and Detection of Hepatitis C Infection
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 142; DB 5;
Pred. No. 1.8e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC -DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: **1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1151-4101PC1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4 PKPQKKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Morgan & Finnegan, L.L.P
                                                                                                                                                                                                                                                                                UMBER: PCT/US94/07088
22-JUNE-1994
                                                                                                        COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: PCT/US95/13660
FILING DATE: 23 October 1995
                                                                                                                                                                                                                                                                                                   FILING DATE: 22-JUNE-1994
CLASSIFICATION
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/083,947
FILING DATE: 28-JUNE-1993
ATTONNEY/AGENT INFORMATION:
NAME: LIN, MARIA C.H.
REGISTAATION NUMBER: 29323
REFERENCE/DOCKET NUMBER: 1151-41(
TELECOMMUNICATION INFORMATION:
TELECHHONE: (212)758-4800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/333,573
FILING DATE: 01 NOVEMBER 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE: Morgan & Fins
STREET: 345 Park Avenue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match
Best Local Similarity 92.9*
....hes 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ) MOLECULE TYPE: peptide
PCT-US94-07088-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CITY: New York
STATE: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PCT-US95-13660-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
```

ò g

1 PKPORKPNRNTNRRPODVKFPGGGOIVG 28

ò

```
RESULT 15
US-08-635-886C-198
US-08-635-886C-198
Sequence 198 Application US/08635886C
Patent No. 655314
Sequence 198 Application US/08635886C
Patent No. 655314
SERRAL INFORMATION
PAPLICANT: DELEYS, Robert
APPLICANT: DELEYS, Robert
TITLE OF INVENTION: INGUINDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C.
TITLE OF INVENTION: VARUES
FILE REPERBNCE: 2752-18
CURRENT APPLICATION NUMBER: US/08/635,886C
CURRENT APPLICATION NUMBER: DE 1940-4025
FRIOR RILING DATE: 1940-40-28
FRIOR APPLICATION NUMBER: EP 93402718.6
FRIOR FILING DATE: 1991-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PREFENTION OF SEQ ID NOS: 286
S
```

Search completed: August 7, 2003, 11:23:50 Job time : 10.6364 secs

us-09-491-146a-26.rapb

```
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                        (without alignments)
231.506 Million cell updates/sec
                                                                                                                                                                      August 7, 2003, 11:20:11; Search time 14.3636 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         // Cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/PCO_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /cgn2_6/prodata/1/pubpaa/US09B_UUBCOMB.pep:*/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*/cgn2_6/ptodata/1/pubpaa/US09_NRW_PUB.pep:*/cgn2_6/ptodata/1/pubpaa/US10B_URRW_PUB.pep:*/cgn2_6/ptodata/1/pubpaa/US10B_UUBCOMB.pep:*/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*/cgn2_6/ptodata/1/pubpaa/US10C_NUBCOMB.pep:*/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                      156
1 PKPORKPNRNTNRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   451899 seqs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                       OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                          US-09-491-146A-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16:
17:
18:
                                                                                                                                                                                                                                                                                                             Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                    Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Searched:
                                                                                                                                                                                               Run on:
```

## SUMMARIES

	Description	Sequence 1, Appli	Sequence 10, Appl	₹	Sequence 8, Appli	11'	14,	78,	46,		152,	152,	2,	4,	42,	44,
	ID	US-10-367-677-1	US-09-851-138-10	US-09-758-308-1	US-09-756-875-8	US-09-921-397-77	US-09-851-138-14	US-09-921-397-78	US-09-851-138-46	US-09-851-138-60	US-09-899-046-152	US-09-878-281-152	US-09-899-046-42	US-09-899-046-44	US-09-878-281-42	US-09-878-281-44
	DB	15	10	6	10	10	10	10	10	10	11	1	11	11	1	11
	Query Match Length DB	44	74	91	97	103	108	113	137	138	166	166	169	169	169	169
æ	Query Match	91.0	91.0	91.0	91.0	91.0	91.0	91.0	0.16	91.0	91.0	0.16	0.16	0.16	0.16	91.0
	Score	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142
	Result No.	-	2	· ~	4		, vc	2	- a	0 0	, 6	- 1	12	7.	7 -	15

Length 44;

DB 15;

Score 142;

91.0%;

Query Match

Sequence 2, Appli Sequence 2, Appli	Sequence 3, Appli	Sequence 76, Appl		Sequence 18, Appl			Sequence 54, Appl	144,	20,	25	54,	144,	Sequence 50, Appl	Sequence 50, Appl	20,	50,	23,	23,	40	-	Sequence 9, Appli		Sequence 20, Appl	Sequence 3, Appli	ì	20,	N	Sequence 2, Appli
10 US-09-929-955-2 . 14 US-10-104-966-2	19-194-		US-09-851-	10 US-09-851-138-18	11 US-09-899-046-50	11 US-09-899-046-52	9668	11 US-09-899-046-144	11 US-09-878-281-50	9-878-28	-818-60	28		us-	11 US-09-995-808-50	08-96-60-SD	10 US-09-941-611-23		US-1	9 US-09-742-659-4	US-09-952	10 US-09-929-955-1			14 US-10-104-966-1	15 US-10-259-275-20	10 US-09-238-076-2	11 US-09-995-937-2
182	191	318	319	319	319	319	319	319	319	319	319	319	809	608	608	809	2894	2894	2985	3011	3011	3011	3011	3011	3011	3011	3012	3012
91.0	91.0	0.16	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91.0	0.16	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91.0		91.0	91.0	91.0
142	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142		142		142	142	142	142	142
16	, e	0.0	56	2 6	22	3 6	90		26	27	20.	50	£ 6	, ,	1.0	. "	3.6	· 10°	9 (9)	37	. ee	3.6	0.4	4.4	42	4	44	4.5

## ALIGNMENTS

```
Sequence 1, Application US/10367677

Squence 1, Application US/10367677

Squence 1, Application US/10367677

Squence 1, Application No. US20030118604A1

GENERAL INFORMATION:

APPLICANT: JOLEON:

APPLICANT: JOLEON:

APPLICANT: LACOUX, XAVIER

APPLICANT: LACOUX, XAVIER

TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND

TITLE OF INVENTION: TREATING AN HCV INFECTION

FILE OF INVENTION: TREATING AN HCV INFECTION

FILE OF INVENTION: TREATING AND USES FOR DETECTING, PREVENTING AND

FILE REFERENCE: 103959

CURRENT APPLICATION NUMBER: US/10/367,677

CURRENT FILING DATE: 2003-02-19

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/FR98/00442

PRIOR APPLICATION NUMBER: EARLIER PILING DATE: 1998-03-05

NUMBER OF SED ID NOS: 11

SGOTWARE: Patentin Ver. 2.1

FENCENCE 1.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS: Ogata, N. et al. TITLE: Nucleotide Sequence and Mutation Rate of the H Strain TITLE: of Hepatitis Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL: Proc. Natl. Acad. Sci.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PAGES: 3392-3396
DATE: 1991
RELEVANT RESIDUES: 2 TO 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 44
```

ô

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                 Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 97;
                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 142; DB 10;
Pred. No. 5.7e-12;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: PIKE, IAN
TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                           Score 142; DB 9;
Pred. No. 5.3e-12;
0; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ...UNKESSE: SUITE 701-E COlumbia Square STREET: 555 13th Street, N. W. CITY: Washington STATE: D. C. STATE: D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPQRKTKRNINRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPORKPNRNTNRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PCT/GB93/00410
                                           60/092,339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-ANG-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTORNEY AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 8, Application US/09756875 Patent No. US20020150990A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-non-
            2001-01-10
                                                                                                                                Patentin version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: ERNST, BARBARA G. REGISTRATION NUMBER: 30,377 REFERENCE/DOCKET NUMBER: 181 TELECOMMUNICATION INFORMATION:
                                                                 1999-07-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
      CURRENT FILING DATE: 2001-01
PRIOR APPLICATION NUMBER: 6
PRIOR FILING DATE: 1999-07-
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin version
SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (202)783-6040
                                                                                                                                                                                                                    ; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity 92.9 Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     D. C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
TITLE OF INVENTION: PROTEINS
                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                        Gaps
                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
   Pred. No. 2.5e-12;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
91.0%; Score 142; DB 10;
Best Local Similarity 92.9%; Pred. No. 4.3e-12;
Matches 26; Conservative 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 10:
                                                                                       1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                            1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
MOLECULE TYPE: peptide
;
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILE REFERENCE: 14114.0349U2
CURRENT APPLICATION NUMBER: US/09/758,308
                                                                                                                                                                                                                                                                   US-09-851-138-10
; Sequence 10, Application US/09851138
; Publication No. US20020183508A1
; GENETAL INFORMATION:
APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 1, Application US/09758308
Patent No. US20020090607A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 74 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AGENTS
92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CITY: HOUSTON STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
US-09-758-308-1
                                                                                             à
                                                                                                                                               쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                          polypeptides selected from
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 113;
                                                            Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 142; DB 10;
Pred. No. 6.7e-12;
0; Mismatches 2;
                                                         Score 142; DB 10;
Pred. No. 6.4e-12;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                    Sequence 78, Application US/09921397

Patent No. US20020151484A1

GENERAL INFORMATION:

APPLICANT: HYBRIGENICS

TITLE OF INVENTION: Dathogenic strain of the TITLE OF INVENTION: pathogenic strain of the TITLE REPERENCE: B4809A - JAZ

CURRENT APPLICATION NUMBER: US/09/921,397

CURRENT FILING DATE: 2001-08-02

PRIOR FILING DATE: 2000-08-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,136
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
; SEQUENCE DESCRIPTION: SEQ ID NO: 14: US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 46, Application US/09851138; Publication No. US20020183508A1; GENERAL INFORMATION: APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       91.0%;
ilarity 92.9%;
Conservative
                                                                91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 78
                                                                Query Match
Best Local Similarity 92.93
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-921-397-78
                                                                                                                                                                                                                                                                         US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                   ò
                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STUVVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
          US-09-921-397-77

Sequence 77, Application US/09921397

Sequence 77, Application US/09921397

Patent No. US20020151484a1

GENERAL INFORMATION:

APPLICANT: HYBRIGENICS

TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and TITLE OF INVENTION: applications thereof

FILE REFERENCE: B4809A - JAZ

CURRENT APPLICATION NUMBER: US/09/921,397

CURRENT FILING DATE: 2001-08-02

PRIOR APPLICATION NUMBER: EP 00402225.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 142; DB 10;
Pred. No. 6.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 14, Application US/09851138 publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 108 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Hepatitis C virus 9-921-397-77
                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Ver. 2.1 SEQ ID NO 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: USA
ZIP: 77210-4433
                                                                                                                                                                                                                                                                                    NUMBER OF SEQ ID NOS: 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -09-851-138-14
                                                                                                                                                                                                                                                                                                                                                  LENGTH: 103
                                                                                                                                                                                                                                                                                                                                                                          TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
```

```
New sequences of hepatitis C virus
genotypes for diagnosis, prophylaxis and therapy.
270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 166;
              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT ARELICATION DATA:
APPLICATION NUMBER: US/09/878, 281
          .
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 142; DB 11;
Pred. No. le-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 11;
          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                              1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                 5 PKPORKTKRNTNRRPODVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91.0%; Score 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                    ; Sequence 152, Application US/09899046; Publication No. US20030008274A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                          PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 152, Application US/09878281 Publication No. US20030032005A1 GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
          ö
                                                                                                                                                                                                                                                                                                                                                                   IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO: 152: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              91.0%;
                                                                                                                                                                                                                                                                                                                                                 Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 92.97
Matches .26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         166 amino acids
        Conservative
                                                                                                                                                                                                                                                               TITLE OF INVENTION: New TITLE OF INVENTION: GENC NUMBER OF SEQUENCES: 270 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy di
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     inear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                      OPERATING SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; MOLECULE TYPE:
US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
                                                                                                                                                                 US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-878-281-152
        26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
        Matches
                                                ó
                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                 ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CURRENTING SYSTEM: PC-DOS/MS-DOS
SOFFWARE: MICROSOFT WORD 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: BP 94870166.9
FILING DATE: CURROWN:
APPLICATION NUMBER: BP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 22 Jun 1995
                                                                                                                                                                                                                                                                                                                          Length 137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 138;
                                                                                                                                                                                                                                                                                                                     Score 142; DB 10; Length 1
Pred. No. 8.2e-12;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 142; DB 10;
Pred. No. 8.3e-12;
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                           NAME: KAMMERER, FATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/COCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHRRACTERISTICS:
LENGTH: 137 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 60:
                                                                                                                                                                                                    TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPORKPNRNTNRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 60, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: USA
21P: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STUYVER, LIEVEN
                                                                                                                                                                                                                                                                                                                   91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-851-138-60
                                                                                                                                                                                                                                                                                                                       Query Match
```

```
GENERAL INFORMATION:
APPLICANT:
APPLICANT:
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
TITLE OF SEQUENCES: 270
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                            New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 169;
Score 142; DB 11; Length 169;
Pred. No. 1e-11;
                                                                                                                                                                                                       Sequence 42, Application US/09878281

Sequence 42, Application US/09878281

Sequence 42, Application No. US20030032005A1

SENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Genctypes for diagnosis, prophylaxis a TITLE OF INVENTION: genctypes for diagnosis, prophylaxis a VINBER OF SEQUENCES: 270

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
COMPUTER: IBM PC Compatible
COMPUTER: DATENTIN SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
CURRENT APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. 1e-1
0; Mismatches
                                              0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 142;
                                                                                 1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                       5 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-878-281-44; Sequence 44, Application US/09878281; Publication No. US20030032005A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
        Query Match 91.0%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : 169 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          , MOLECULE TYPE: protein US-09-878-281-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        , MOLECULE TYPE: protein US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                           US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                             à
                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TILE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy
NUMBER OF SEQUENCES: 270
NUMBER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                          New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy, 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 11; Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                          Indels
                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylax NUMBER OF SEQUENCES: 270
COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: DatentIn Release #1.0, version #1.25 (EPO)
CURRENT APPLICATION DATA:
                                          5
                          Pred. No. 1e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91.0%; Score 142; DE
92.9%; Pred. No. 1e-1
Live 0; Mismatches
                                              0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                               1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                     APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 44, Application US/09899046 Jublication No. US20030008274A1
                                                                                                                                                                                                                   Sequence 42, Application US/09899046 Publication No. US20030008274Al GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
                        Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 : 169 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 91.0
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            , MOLECULE TYPE: protein US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 13
US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-899-046-42
                                                                                                                                                                                           RESULT 12
US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ф
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                             õ
                                                                                                                              g
```

Ţ

Gaps ö Ouery Match 91.0%; Score 142; DB 11; Length 169; Best Local Similarity 92.9%; Pred. No. 1e-11; Matches 26; Conservative 0; Mismatches 2; Indels (

. O

g ò

Search completed: August 7, 2003, 12:01:13 Job time : 15.3636 secs